Abstract

Title of thesis: THE RELATIONSHIP BETWEEN THE LEVEL OF IMPLEMENTATION OF INSTRUCTIONAL CONSULTATION TEAMS AND STUDENT GOAL ATTAINMENT,

Allison Marie LaFleur, Master of Arts, 2006

Thesis directed by: Professor Sylvia Rosenfield
Department of Counseling and Personnel Services

A number of prereferral and problem-solving teams have emerged since PL 94-142 to provide consultation and intervention planning for teachers with difficult to teach students. Instructional Consultation Teams (IC Teams) is one model that attempts to improve student academic and behavioral performance through the use of a structured problem-solving process. This study examined the archived data of 417 IC Team consultant-teacher dyads to determine the relationship between implementation of the collaborative process and student goal attainment.

High levels of goal attainment were assessed for cases that provided full SDF documentation. Implementation in the dimensions of Clear Communication, Intervention Implementation and Intervention Evaluation related to a small degree with student goal attainment. However, 42% of cases did not provide sufficient information to determine goal attainment. Those cases that did fully document SDF data had higher implementation scores than did those that provided partial or insufficient documentation of critical SDF components.

THE RELATIONSHIP BETWEEN THE LEVEL OF IMPLEMENTATION OF INSTRUCTIONAL CONSULTATION TEAMS AND STUDENT GOAL ATTAINMENT

by

Allison Marie LaFleur

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Arts
2006

Advisory Committee:

Sylvia Rosenfield, Ph.D., Chair
Todd A. Gravois, Ph.D.
William Strein, Ed.D.

Table of Contents

List of Tables

Chapter 1

Introduction

Since the authorization of the special education law PL 94-142 in 1975, school systems have recognized the need to improve educational services to accommodate more students in the general education setting prior to consideration of special education (Villa & Thousand, 2003). This movement has resulted in the development of several prereferral team models designed to serve as problem-solving units to assist teachers and other school professionals in improving services for students in the general education environment. For the most part, prereferral teams have functioned as indirect service delivery or consultation systems in which school professionals work collaboratively with one another to better serve a third-party client, namely a student, in the general education environment (Fuchs et al., 1990; Nelson & Smith, 1991; Welch, Brownell, & Sheridan, 1999).

Teaming and Consultation Models

Teaming models have emerged in various forms over the years, differing in membership and process; however, for the most part their goals are similar in that each team strives to indirectly assist students having difficulty in regular education. Teacher Assistance Teams (TATs) emerged early and were comprised mainly of general education teachers whose purpose was to assist other educators (Chalfant, Pysh, & Moultrie, 1979). The emergence of the Prereferral Intervention (PI) model (Graden, Casey, & Christenson, 1985) did not specify team membership, but stressed that a collaborative consultative relationship be in place between two individuals. In addition, the PI model initiated a movement in which consultation was seen as a method to prevent

inappropriate referrals to special education. Other teams have utilized the services of administrative personnel and specialists trained in a specific aspect of the prereferral process (Rosenfield & Gravois, 1999). Mainstream Assistance Teams (MATs) were designed specifically to provide assistance to teachers for problematic student behaviors through dyadic relationships in which the consultant was trained in the methods of behavioral consultation (Fuchs, Fuchs, Gilman, et al., 1990). Teams also differed in whether a group of individuals functions as the main agent of problem-solving (Chalfant et al., 1979) or whether problem solving occurs directly between an individual consultant and consultee (Rosenfield & Gravois, 1996).

Outcomes of Teams

Despite the abundance of teams that have emerged since PL 94-142, there is some controversy regarding the effectiveness of any given model in producing beneficial outcomes for students (Nelson & Smith, 1991; Sheridan & Welch, 1996; Welch et al., 1999). Reduced special education referral rates have been one of the predominantly cited outcomes arguing for the effectiveness of teams, with subjective measures of teacher satisfaction and positive staff attitudes also commonly noted as merits of prereferral interventions (Chalfant & Pysh, 1989; Fuchs, Fuchs, & Bahr, 1990; Fuchs, Fuchs, Gilman, et al., 1990; Graden, Casey, & Bonstrom, 1985). However, much additional research on team effectiveness is needed as some studies produce contradictory evidence about the outcomes even within a given model (Chalfant & Pysh, 1989; Fuchs, Fuchs, & Bahr, 1990; Graden, Casey, & Bonstrom, 1985; Nelson & Smith, 1991).

In addition, only a small number of studies have actually measured student behaviors in the classroom to provide evidence of effectiveness (Fuchs, Fuchs, & Bahr,

1990; Fuchs, Fuchs, Gilman, et al., 1990), and even fewer provide information about the students' academic functioning as a result of prereferral interventions (Chalfant & Pysh, 1989). Thus, despite the numerous publications on prereferral teams, findings and measurement techniques are still inconsistent and few directly measure the impact that teaming interventions have on actual observed student outcomes (Nelson & Smith, 1991; Welch et al., 1999).

Goal Attainment Scaling

Goal Attainment Scaling (GAS) is one method for recording both observed and indirect student outcomes. The GAS process has been identified for use in monitoring individual student achievement in a variety of behavioral or academic areas and includes goal setting, specification of programs and time for evaluation, goal attainment scaling or ranking, and evaluation of goal attainment (Kiresuk, Smith, & Cardillo, 1994). GAS has been used to evaluate the effectiveness of interventions for special education populations including students with autism, intellectual disabilities, learning disabilities, and multiple-handicaps, in addition to students at risk for academic difficulties (Glover, Burns, & Stanley, 1994; Kratchowill, Elliott, & Bussee, 1995; Oren & Ogletree, 2000; Wehmeyer, Hughs, Agran, Garner, & Yeager, 2004). In addition, it has been cited as useful for organizing the direction of services and involving parents and students in the setting of goals (Oren & Ogletree, 2000). While GAS does not include a standardized measure by which students can be compared, it provides flexibility in that it allows for outcome measurement for students with a variety of needs and educational goals (Maher, 1983).

Level of Implementation

Level of implementation has been identified as the extent to which a program or initiative is operationalized as planned. Level of implementation or treatment integrity is often measured by the presence or absence of critical components during the implementation of any given model (Fudell, 1992). It has been recognized that evaluation of teaming models cannot occur without evidence that the crucial components of a model are in place and that other conflicting elements are not interfering with implementation. Without awareness as to the fidelity of prereferral treatment efforts, it is impossible to tell whether failure of a model is due to the model itself or merely to insufficient implementation (Gresham, 1989). Another difficulty in assessing the effectiveness of teams is that few studies include measurement of the actual implementation integrity with which prereferral models are being implemented (Nelson & Smith, 1991). Some studies have attempted to address treatment integrity or level of implementation in consultation and prereferral models; however, findings are inconsistent and the measures used to determine integrity vary widely (Bahr, Whitten, Dieker, Kocarek, & Manson, 1999; Flugum & Reschly, 1994; Gresham & Gansle, 1993; Kovaleski, Gickling, Morrow, & Swank, 1999; Wickstrom, Jones, LaFleur, Witt, 1998; Witt, Gresham, & Noell, 1996).

While overall it appears that implementation integrity makes a difference for prereferral outcomes, further research is needed to determine how actual elements of the process lead to gains in student behavior and academics (Bergan & Tombari, 1976; Fudell, 1992; Telzrow, McNamara, & Hollinger, 2000). In addition, there is a need to

ensure that measures of treatment integrity are objective and reliable as self-reports may be inconsistent (Robbins & Gutkin, 1994).

Instructional Consultation Team Model

The Instructional Consultation Team (IC Teams) Model (Rosenfield & Gravois, 1996) is one prereferral intervention model that focuses on assisting teachers with students' academic and behavioral concerns through one-on-one case consultation and a focus on providing an instructional match between the student and the environment. Additional components of the model include delineation of precise steps to be followed throughout the collaborative consultation process and training in effective communication strategies. While research has validated this model as beneficial in reducing special education referrals and placements (Gravois & Rosenfield, 2002), improving teacher skill, confidence and satisfaction (Costas, 2003) and reducing disproportionate referrals of minority students (Gravois & Rosenfield, 2002), there remains a need to link implementation of the model to student outcomes (Costas, 2003; Levinsohn, 2000; Weiner, 2002).

Statement of the Problem

Despite a plethora of teaming models in the literature within the past 25 years, differences exist as to how to empirically validate these models. While a number of studies utilize subjective informant reports and school wide referral rates, little research has provided information on the implementation integrity of these models and its relation to objective student outcomes (Nelson & Smith, 1991; Welch et al., 1999). The IC Team (Rosenfield & Gravois, 1996) model has been cited as effective in reducing special education referrals and placements, improving teacher skill, and promoting beneficial

student outcomes; however little has been done to relate IC Team outcomes to the integrity with which the model is implemented (Costas, 2003; Gravois, 1997; Weiner, 2002). In addition, failure to fully document information about student goal attainment makes it difficult to accurately assess this student outcome (Levinsohn, 2000; Gravois & Rosenfield, 2001). Thus, the purpose of this research was to address the gap in the literature by assessing the implementation level of the IC Team model and its relation to student outcomes as measured by Goal Attainment Scaling. The seven critical dimensions of the IC Team process (LOI-R, Fudell, 1992; Vail, 1996) were examined to determine if particular elements of the process differ for cases at different levels of goal-attainment. In addition, because program evaluations of IC Team implementation indicate that a number of cases fail to document baseline, goal, and intervention data necessary to determine goal attainment, an additional focus of the study was in determining how elements of the process differ in situations where goal attainment can not be evaluated due to incomplete data (Levinsohn, 2000; Gravois & Rosenfield, 2002). The study was guided by these research questions:

1. For IC Teams cases, how well are critical SDF components documented?
2. For cases that have not fully documented critical SDF components, what components are missing from the documentation?
3. Do cases with full, partial and insufficient documentation of critical SDF components have different scores on the Level of Implementation- Revised Collaborative Process dimensions?
4. For IC Team cases providing full documentation of critical SDF components, what are the levels of goal attainment?

5. Do different levels of treatment integrity of the IC Team model, as measured by the Level of Implementation Collaborative Process scores, relate to different goal attainment rates, when critical components are fully documented?

Definition of Terms

Instructional Consultation Team (IC Team) Model. The IC Team Model is a prereferral problem-solving team model that provides assistance to teachers in creating an instructional match between the student, task and environment to better accommodate students with diverse needs in regular education. The model is based on a five stage problem-solving process including the following stages:

1) Contracting: explanation of the process and agreement to work together on the problem stated,
2) Problem-Identification: specification of the problem in observable terms and assessment of student's current functioning,
3) Intervention Design: formulation of an agree-upon intervention plan,
4) Intervention Implementation: execution of intervention plan, and
5) Intervention Evaluation: monitoring data collected during intervention to determine effectiveness or need for modification of the intervention (Rosenfield & Gravois, 1996).

Goal Attainment Scaling (GAS). Goal Attainment Scaling (Kiresuk et al., 1994) is a means of measuring student achievement in academic or behavioral domains that allows for individual goal setting based on a student's current level of functioning and the desired or expected level of functioning for success in the classroom. It involves data

collection before, during and after an intervention to determine if progress is being made according to the levels of behavior and time line specified. Goal attainment is documented by assessing the student's rate of progress over at least three data points, comparing the data to the baseline performance and to the goals set and determining whether the performance is below baseline, consistent with baseline, toward the goal, or has met a short term, interim or long term goal. These descriptors are then converted to a continuous numeric scale with the midpoint representing the expected level of goal attainment (Kiresuk et al., 1994).

Level of Implementation (LOI)/ Treatment integrity. LOI is a measurement of the extent to which a model or intervention is implemented as planned by checking for the presence or absence of certain key indicators of the model's process. Level of implementation is assessed within the IC Team process through the Level of Implementation-Revised measure, which engages the teacher and case manager for each case in a semi-structured interview. Scores are assigned for each case based on the presence or absence of certain indicators critical to the IC Team Model. Certain items also require that the case manager and teacher provide similar responses or agree on how the model was implemented (Fudell, 1992; Vail, 1996). Components of LOI include measurement of the Collaborative Process dimension, which describes the degree to which various critical components of the problem-solving process were implemented, and measurement of the Team Process. Variables within the Collaborative Process dimension include clear communication, contracting, problem identification, intervention recommendation, intervention implementation, intervention evaluation, and curriculum based assessment.

Consultant/Case Manager. A consultant is an individual providing assistance to a student indirectly by engaging in problem-solving and skill development activities with the student's teacher. In the IC Team model, the consultant is referred to as the case manager.

Consultee. A consultee is the teacher or professional requesting assistance for problem-solving or skill development support from a consultant in working with a difficult to teach student or group of students.

Chapter 2

Review of the Literature

The purpose of this chapter is to review the research relevant to this study of treatment integrity and student outcomes of Instructional Consultation Teams (IC Teams). This chapter will review the theoretical background and structure of teaming models, selected research on outcomes of teaming models, a description of the theory and research behind IC Teams, design and findings related to student goal attainment scaling (GAS), and considerations for evaluating implementation integrity of teams.

Overview of Teams

Teaming models began to emerge in the late 1970's and early 1980's with the authorization of P.L. 94-142. The law indicated that prior to consideration for special education services, the child had to be assisted by other documented support methods and interventions (Villa & Thousand, 2003). In addition, Public Law 94-142 mandated that multidisciplinary teams (MDTs) be used to make decisions regarding special education referral and placement (Rosenfield & Gravois, 1999). Thus, various teaming models began to emerge as a means of facilitating delivery of services to students experiencing difficulty and assuring that decisions were made according to multiple professionals' combined expertise (Nelson & Smith, 1991; Welch et al., 1999).

Early teaming models. The development of indirect service teams was primarily influenced by two other models for psychological and educational intervention. Collaborative Consultation, which stemmed from the development of mental health services in the 1950's, provided a framework for using indirect assistance to a student through a teacher's work with a specialist (Sindaler, Griffin, Smith & Watanabe, 1992).

Similarly, Teacher Assistance Teams emerged in the 1970's as problem-solving groups where teachers could assist each other in serving difficult to teach students (Chalfant et al., 1979).

Teacher Assistance Teams (TATs) were developed for the purpose of providing teachers with a systematic means for consulting with one another to problem-solve and develop interventions for use with students of varying abilities in general education classrooms. The assumptions underlying TATs included the beliefs that teachers in regular classroom environments can assist students with academic and behavioral problems and, while teachers may need assistance in accomplishing this task, they have considerable knowledge and talent in teaching and helping one another to provide services. In addition, TATs were founded on the principle that they can help resolve more problems and help students more effectively when working together, as opposed to working alone. Typical teams consisted of three grade-level educators, although some teams incorporated multidisciplinary membership, gathering to assist a fourth referring teacher with a specific concern. TATs main goals were to provide a support system that could address a variety of concerns while empowering classroom teachers and minimizing problem analysis from a child-perspective. In addition, TATs were designed to support and complement both general and special education classrooms instead of replacing special education (Chalfant et al., 1979; Chalfant & Pysh, 1989).

In 1985, Graden, Casey and Christenson proposed an alternative model utilizing the similar notion of providing consultative services to teachers instead of direct service to students. They outlined the Prereferral Intervention (PI) Model to clearly define what processes and functions comprise successful consultative services. The PI Model

proposed that special education is not necessarily the most appropriate way to serve many mildly handicapped or difficult to teach students and was a response to the growing number of students referred to and placed in special education services. Instead of providing direct service to students, resources are to be re-directed through regular education teachers to serve more students in the general education environment. The focus is on students' learning in an ecological context instead of on problems viewed as internal to the student. Graden, Casey, and Christenson outlined the stages fundamental to a prereferral intervention system to include: 1) a request for consultation, 2) consultation involving a collaborative relationship, definition of concerns in objective and behavioral terms, problem-solving steps, intervention implementation and evaluation, 3) observation and collection of behavioral and environmental information, 4) conferencing to share information and make decisions with a "Child Review Team," 5) a formal referral for evaluation if appropriate, and 6) a formal program meeting to determine the most appropriate services.

In addition, assumptions of this consultative, prereferral system include shared power and collaborative decision-making made throughout the process, ownership of the decision to intervene by the teacher, the development of interpersonal and problem-solving skills and related content and systems theory expertise of the consultant. Thus, by providing interventions to assist students in the regular education environment, PI models strive to reduce inappropriate referrals, evaluations and placements in special education and provide appropriate and needed assistance to students and teachers in the least restrictive environment (Graden, Casey, & Christenson, 1985).

Another early prereferral intervention model proposed in the late 1980's incorporates elements from both the teaming and consultation literature. Mainstream Assistance Teams (MATs) were developed under the least restrictive doctrine set by PL 94-142 and focus on providing preventative assistance to teachers to increase their skill in handling future behavioral problems while reducing special education referrals. A significant component of MAT service delivery is that it utilizes the expertise of special service personnel who are trained in the process of behavioral consultation. In addition, MATs adopt the structured problem-solving stages put forth in the consultation literature and stress the importance of collecting data, setting goals and evaluating goal-attainment continually throughout the process (Fuchs, Fuchs, Gilman, et al., 1990).

Team composition. Initially, prereferral teams were comprised mainly of educators, as in the TAT model, but more recent models have recognized the benefit of having various specialists, special educators, and administrators serve in consultative roles (Fuchs, Fuchs, Gilman, et al., 1990; Rosenfield & Gravois, 1996, 1999). However, the foremost issue in considering membership for teams is that those involved include: members with expertise needed to make the best decisions, those affected by decisions, and those with interest in participating (Thousand & Villa, 1992). While the membership and procedures of various prereferral intervention models differ, they all share some common characteristics in that they attempt to accommodate students' needs in the least restrictive environment possible, focus on prevention such that students are assisted before special services are needed, and include membership of one or several individuals serving in a consultative role to support the student indirectly through the teacher (Fuchs, Fuchs, Gilman, et al., 1990).

Elements of collaborative teams. Some aspects of successful collaborative teams have been identified by Thousand and Villa (1992), who define collaborative teams as groups of people who coordinate their work towards achievement of one or more commonly agreed-upon goals, possess of a belief system that all members of the team have unique expertise, value each member's input equally, distribute leadership tasks and functions, and employ a collaborative teaming process. They qualify an effective collaborative teaming process as one in which members have frequent face-to-face interaction, exemplify feelings of positive interdependence, work to develop group interpersonal skills, assess and discuss team functioning, goal setting and relationships, and employ methods for holding one another accountable for responsibilities and commitments. While much is written on the theory of effective teams, there is a greater need for empirical validation of the aspects of different teaming processes.

While recognized for their potential to use differing perspectives to increase the quality and validity of solutions generated by the team, the value of MDTs was based primarily on intuitive reasoning and not research (Rainforth & York-Barr, 1997; Rosenfield & Gravois, 1999). In addition, while the law mandated MDTs, it gave little guidance in terms of the processes these teams should carry out in order to be successful. Criticisms of early MDT's included a lack of systematic approaches to collect and analyze data, minimal parent or educator participation, unstructured decision-making processes, ambiguous role definitions, low accountability and a lack of collaboration. Thus, while the literature is teeming with different models and suggestions for the implementation of various teams, few are consistently backed by research (Rosenfield & Gravois, 1999).

Effectiveness of teams

With the development of several models of problem-solving and prereferral intervention teams, there is a need to determine uniform methods for evaluating the effectiveness of various models. Sindaler et al. (1992) identify five measures typically cited in the literature as indicators of effective prereferral practice: reduction in referrals to special education, improved student performance on academic or behavioral concerns, altered teacher expectations or perceptions, teacher/parent/student satisfaction and an improvement in educational practice or teacher skill. However, research on the actual effectiveness of prereferral interventions in producing these outcomes is relatively inconsistent and many studies vary as to how to define and measure the desired outcomes (Chalfant & Pysh, 1989; Graden, Casey & Bonstrom, 1985; Nelson & Smith, 1991; Sheridan & Welch, 1996; Welch et al., 1999).

Referral and placement. One method by which prereferral systems are frequently evaluated is by referral and placement rates before and after the intervention. Early investigation on the effectiveness of the PI Model by Graden, Casey and Bonstrom (1985) indicated that in four of six schools examined, special education referral and placement rates decreased after a year of utilizing a prereferral intervention system. However, in one school, there was no difference in referral or placement rates in the subsequent year, and a final school actually experienced an increase in placements. Several studies on the TAT model found that implementation of the model resulted in appropriate referrals to special education in 78-100% of the cases, with a 63% drop in inappropriate referrals compared to the pre-implementation year (Chalfant & Pysh, 1989). Similarly, MAT research has demonstrated special education referral rates of 8%

and 13% for short and long implementation conditions, compared to a 50% referral rate for students from a control group. However, this study tracked only 24 students in each of the implementation conditions and 12 students in the control group (Fuchs, Fuchs, & Bahr, 1990). However, in only one study of TATs was appropriateness of referrals considered despite the notion that reduction of inappropriate referrals is one of the fundamental reasons for implementing prereferral teams (Chalfant & Pysh, 1989).

Observable behavior change. A second means of determining prereferral effectiveness is by tracking observable problem behaviors before and after the intervention has been implemented. In a study of short and long MAT implementation conditions, the observed frequency of target behaviors decreased so extremely that at the end of the interventions, the difficult to teach student group was indistinguishable from the control group with respect to frequency of problem behaviors (Fuchs, Fuchs, & Bahr, 1990). Similarly, another study of Behavioral Consultation (BC) outcomes reflected between 63% and 88% of students demonstrating reduced problem behaviors following intervention, in comparison to 29% of control students. Despite this convincing data, the studies measuring this outcome are sparse and are generally focus on measurement of behavior as opposed to academic skills (Nelson & Smith, 1991; Welch et al., 1999).

Surveys and subjective ratings. While special education rates and behavioral observations are relatively objective measures of prereferral effectiveness, a majority of models continue to emphasize subjective reports and ratings about improvement as primary outcomes. Research on TATs measured student behavioral functioning, team effectiveness, progress towards goal attainment and resolution of problem behaviors all through subjective reports and ratings. One study judged the team's overall effectiveness

and attributed their success to variables of principal support, team attitudes and training, team performance and faculty support/participation all through subjective reports (Chalfant & Pysh, 1989).

The PI system was also supported by satisfaction ratings of classroom teachers and principals who believed the process to be helpful and the interventions effective (even when referral rate effectiveness was not observed). In addition, principals believed that PIs improved teacher tolerance and competence in working with diverse students. To explain some of the variations in their findings, Graden, Casey, and Bonstrom (1985) commented that schools with successful interventions were those that had the support of at least one individual claiming consultation as her primary role.

MAT research also claims effectiveness through largely subjective reports. One study found that PI techniques resulted in fewer problem behaviors of difficult to teach students, increasingly positive teacher attitudes, and lower special education referral rates when compared to a control group. However, the cases in this study benefited from at least weekly support from the research team, and it is unclear whether the findings would be replicated if this support was not present (Fuchs, Fuchs, Gilman et al., 1990). In a second study of both long and short MAT implementation conditions, teachers in both implementation conditions responded positively to questions about intervention feasibility and effectiveness ratings. In addition teacher perceptions of the students' behavior were more positive in short implementation conditions with regard to attention problems and anxiety/withdrawal problems from pre to post intervention (Fuchs, Fuchs, & Bahr, 1990).

While support for prereferral teams seems positive at first glance, a large amount of research still defines effectiveness in terms of subjective reports of the professionals involved, and few models demonstrate consistent, measurable impact. A review by Nelson and Smith (1991) examined outcomes of six different prereferral teams and found that the majority of articles indicated success in reducing referral rates to special education by about 40-50%. In addition, the authors reported positive findings in terms of teacher satisfaction, attitudes towards assisting students with differential abilities and perceived improvements in student behavior. However, only one study on Teacher Resource Teams reported effectiveness in objective terms of student goal attainment (Maher, 1986 as cited in Nelson & Smith, 1991). Other reviews (Sheridan & Welch, 1996; Welch et al., 1999) of prereferral and consultation systems cite positive outcomes, but also highlight questionable methodology and subjectivity of outcome measures as major shortcomings of research to date. In addition, as many prereferral models focus primarily or solely on behavioral consultation (Fuchs, Fuchs, Gilman, et al., 1990; Sheridan & Welch, 1996), there is limited, if any, evidence whether students' outcomes under these models reflect any academic progress.

Thus, prereferral intervention and teaming research generally indicates effectiveness in reducing special education referrals and improving teacher satisfaction ratings. Although teacher and team member reports are important measures of satisfaction for those involved in the team, more direct and objective measures of student outcomes are needed to determine actual effectiveness of these models. Few studies report information on goal attainment or other observable measures of prereferral effectiveness. In addition, teams that report focusing on Behavioral Consultation only or

deal mostly with behavioral requests are unlikely to assist students that do have a real difference in instructional needs. Instructional Consultation Teams (Rosenfield & Gravois, 1996) is a prereferral model that focuses on providing instructional assistance to teachers and measuring outcomes in direct, observable terms.

Instructional Consultation Teams (IC Teams)

The Instructional Consultation Team model (Rosenfield & Gravois, 1996) was developed on the premises of instructional consultation (Rosenfield, 1987), which examines the match between the student, task and instructional strategies when trying to address learning obstacles in education. The key focus of Instructional Consultation Teams is to work with teachers and school professionals to provide the best instructional match for each student's learning needs (Rosenfield, 1987). Thus, the framework for the IC Team model is based on the collaborative, problem-solving relationship between the case consultant and the teacher and the adoption of three primary assumptions that 1) all students are learners, 2) the focus should be on instructional match, not place, and 3) that problem-solving communities in schools set a context for effective academic and behavioral development (Rosenfield & Gravois, 1996).

Other key components of the IC Team system involve a non-hierarchical relationship among professionals, training in basic listening and clear communication skills, and extensive evaluation of the IC Team components. In addition to the basic assumptions and framework of IC Teams, the collaborative process is guided by five distinct stages of consultation: entry and contracting, problem identification and analysis, intervention planning, intervention implementation, and resolution/termination. The progression of each case through these five steps is initiated by a teacher's request for

assistance and monitored by the IC Team and systems manager. In addition, if any individual teacher-consultant pairing reaches an impasse in their case, the larger IC Team serves to assist in the problem-solving process (Rosenfield & Gravois, 1996).

Instructional Consultation Team Outcomes

The IC Team model has been evaluated in terms of student, teacher, team and school-wide outcomes (Costas, 2003; Gravois, 1996, 1997; Gravois & Rosenfield, 2002; Levinsohn, 2000; Weiner, 2002). A primary focus of IC Teams outcome research is the reduction of inappropriate referrals to special education. Data on school-wide referral rates of IC Team schools consistently demonstrate decreases in inappropriate referrals to special education once the IC Team model is in place (Gravois, 1996, 1997; Gravois & Rosenfield, 2002). In comparison with schools implementing Student Support Teams (SSTs), an alternative prereferral intervention model, IC Team students were less likely to be screened inappropriately for special education and were less likely to be placed in special education (Levinsohn, 2000). Similarly when comparing IC Teams to IEP teams, Weiner (2002) found that a lower percentage of referred cases were evaluated in IC Teams, resulting in fewer placements of the students originally referred to IC Teams. In addition, IC Teams have been demonstrated to result in more appropriate referrals and placements of students in traditionally overrepresented ethnic minority populations (Gravois & Rosenfield, 2002; Levinsohn, 2000).

Measurement of teacher satisfaction and skill development has also been a large focus of the research evaluating prereferral teams, as consultation models function under the assumption that assisting the teacher will result in generalization of beneficial practices with other students of concern and with classrooms as a whole (Sheridan &

Welch, 1996). Costas (2003) examined IC Team schools and found that 83.5% of teachers reported that they were "satisfied" or "very satisfied" with the process, and 80.6% reported learning more than one new strategy from their experience. In addition, a majority of teachers surveyed reported being "very confident" in handling similar problems in the future. While satisfaction correlated positively with anticipated future use of strategies and perceived outcome gains made by students, it did not relate significantly with actual use of the newly-learned strategies. Thus, these data suggest that teachers may feel positively and respond optimistically about their experience with IC Teams, but evidence is still lacking as to whether satisfaction actually results in improved practice. In addition, a limitation of this study is that it relies on subjective teacher reports of what was learned and what was implemented, instead of objectively measuring teachers' skills.

The IC Team literature has attempted to address the need for evidence in terms of student outcomes, yet findings are inconsistent. Levinsohn (2000) examined the reading progress of IC Team students in comparison with students served through a Student Support Team (SST). While students served through the IC Team process did not have significantly higher reading achievement scores than the comparison group, there was some indication of reading gains. Students referred to IC Team were significantly lower in reading achievement than SST students at the beginning of the school year; however, at the post-test, no differences in reading achievement were measured. While this study presents some evidence for effectiveness of IC Teams in improving students' achievement, there is difficulty in using criterion-based measures to assess the gains of all students. In addition, much more work needs to be done to clarify the conditions and

practices in effect that actually lead to objective student gains. Thus, while the IC Team model has begun to measure student and teacher outcomes in direct, objective ways, much work still remains to be done.

Program Outcome Evaluation

A review by Sheridan & Welch (1996) found that 50% of articles produced on teaming and consultation outcomes assessed effectiveness using multiple outcome measures. However, a majority of these seemed to be self-reported ratings of the process or informant reports, with 49% using satisfaction measures and less than half addressing whether the outcomes of consultation were socially-valid in real life settings. Similarly, Welch et al. (1999) found that a majority of studies on school-based problem solving teams were evaluated only in terms of team member self-reports of outcomes and satisfaction. In addition, Nelson & Smith's (1991) review of studies on prereferral teams indicated that success was generally determined by reductions in referrals to special education, high teacher satisfaction and improved attitudes of teachers towards students with differential abilities. With only one study examining teacher and student achievement on set goals, the authors suggest that further research be conducted to evaluate prereferral teams in terms of student achievement and gains. Although teacher and team-member reports are valuable in that they provide a measure of perceptions and satisfaction of the professionals involved in the process, it is the actual impact of teaming models on student outcomes that is likely to be important in determining whether a child will be successful in the general education setting (Flugum & Reschly, 1994).

While it seems that multiple methods of assessing teaming outcomes have been used, the literature indicates a need for improving research on teams to include long-term,

direct measures of student outcomes that can be generalized to appropriate socially-valid settings. In addition, they recognize the need to specify the extent that the team is being implemented consistently with the model in evaluating success and determine the process variables involved in each team that lead to various outcomes (Nelson & Smith, 1991; Sheridan & Welch, 1996; Welch et al., 1999).

Goal Attainment Scaling

Goal Attainment Scaling (GAS) is a method of setting and evaluating individualized goals as a result of a treatment program. GAS was originally developed as a means for evaluating clinical mental health interventions, but has since been identified as a useful instrument for evaluating progress in various medical, clinical, and educational environments. The GAS process includes the following steps 1) identification of the issues to be the focus of treatment, 2) translation of concerns into specific goals, 3) selection of a title for each concern and prioritizing concerns, 4) selection of a specific behavioral indicator for each concern, 5) specification of the level of behavior expected as an outcome, 6) specification of levels considered to be somewhat more, somewhat less, much more and much less than what is expected in terms of outcomes, and 7) assignment of a goal attainment score based on the level obtained following the expected timeframe of the intervention. Maher (1983) has also proposed procedures for using GAS in educational settings that include the definition of goals in terms of observable and measurable behaviors that relate to intrinsic learning-oriented outcomes, weighing goals in terms of importance, and considering the likelihood of achieving each goal based on information individualized to the student. GAS allows educators to specify goals consistent with the individual needs of each student, the nature

of the program and the policies of the institution. In addition, it provides a flexible yet predetermined schedule for evaluating progress in which various levels and importance of goals can be specified.

Goal Attainment Scaling for individual progress monitoring. GAS methods have been applied to both individual and program monitoring research for many types of populations in mental health, medical and educational settings (Kiresuk et al., 1994). Glover et al. (1994) found GAS to be an effective tool in measuring behavioral progress of learning disabled students in a residential setting, as staff reported the system to be feasible to implement and useful in assisting the planning and achievement of focused, realistic goals. In a general education setting, Wehmeyer et al. (2003) used GAS to measure effectiveness of self-monitoring techniques for adolescents with intellectual disabilities. GAS has also been used to measure parents' perceptions of consultation outcomes in a program evaluation study of communication styles during consultation for disabled and non-disabled students (Grissom et al., 2003). In addition, Oren and Ogletree (2000) reported that GAS measures fit well with practices of task analysis and goal setting in Individualized Education Plans (IEPs) and Individualized Family Service Plans (IFSPs) for students with autism. They used GAS to facilitate input from multiple sources and believed it to be effective in allowing families and service providers to prioritize and weigh goals for special education students. Thus, GAS can be used to measure progress towards an outcome and facilitate change during intervention efforts.

GAS as an effectiveness measure. Goal Attainment Scaling has also been used by various prereferral intervention models to claim effectiveness of treatment interventions. This use of GAS is focused more on evaluating outcomes at the end of an intervention, as

opposed to using data and goals to monitor progress during an intervention. Chalfant and Pysh (1989) measured goal attainment using a five-point Likert scale and found that of students assisted by TATs, 44% made great or considerable gains, 35% made moderate progress and 21% made little or no progress. An examination of MATs revealed that students attained goals set in about 75% consultation cases implementing short and long versions of consultation within guidance counselor/general education teacher pairs (Fuchs, Fuchs, & Bahr, 1990). A limitation of the Fuchs, Fuchs, and Bahr study is that cases benefited from weekly assistance from research staff members, a condition not likely to be implemented in a practical setting. The GAS method was also found to provide evidence of high levels of goal attainment for training programs with high implementation integrity. For example, 80-100% of cases implemented by behavioral consultants and conjoint behavioral consultants demonstrating at least 80% mastery of session objectives were reported to have attained or partially attained behavioral objectives (Kratchowill et al., 1995; Sheridan et al., 2001). Thus, GAS studies have found it to be a sensitive and useful measure for evaluating pre-referral intervention outcomes.

Reliability and validity of GAS. The Goal Attainment Scaling method has been evaluated for interrater reliability. While Kiresuk, Smith, and Cardillo (1994) state that additional research is needed to determine a more specific level of reliability, their review shows that GAS is at least as reliable as other rating scales and generally produces a reliability coefficient of $r = .71$ to $r = .93$. Research on the content validity of GAS has primarily centered on whether goals set are relevant for the treatment plans of individual clients. Preliminary studies suggest that 78% to 97% of goals set were relevant for

clients' concerns and treatment. However, because the GAS process is not specific to any particular concern or restricted to any intervention, measurement of criterion-related validity is difficult. GAS has been found to relate moderately to other outcome measures, suggesting that it assesses some of the same features as other procedures, although GAS's real value is in its ability to assess the change that is produced during treatment procedures. Studies indicated that GAS was a particularly sensitive method of measuring treatment-related change (Kiresuk et al., 1994).

Limitations to GAS. While GAS offers advantages in terms of flexibility of use with several types of interventions (academic, social, emotional and behavioral), there are also limitations to individualized goal scaling. Because goal attainment is individualized, it is difficult to compare outcomes across settings or even within one program. It is possible that goals may be weighted very differently depending on the student or professionals involved, such that attainment for one individual may mean something very different than attainment for another individual (Maher, 1983). Thus, GAS may be most useful for rating individuals' progress in relation to previous performance as opposed to comparing goal attainment across several students.

Goal Attainment and IC Teams

Goal Attainment Scaling has been utilized as a method of measuring student outcomes in the IC Team project, both during intervention efforts and for program evaluation purposes. Thus, goal attainment data are considered to be an indicator that the IC Team process was followed with integrity (process outcomes), but are also used for program evaluation (effectiveness outcomes). However, a lack of consistent data documentation in IC Team cases interferes with the use of GAS to measure viable student

outcomes. High rates of student academic and behavioral goal attainment are evident as outcomes of the IC Team process in that goals set are reached in approximately 80% of IC Team cases (Gravois & Rosenfield, 2002). However, due to poor documentation in many IC Team cases, findings on goal attainment are difficult to interpret. For instance, Moniodis (1996) examined form completion in IC Team cases and revealed that 72% of cases failed to correctly document and graph the student's baseline performance and 75% had incomplete documentation of goal statements.

Similarly, Levinsohn (2000) observed goal attainment rates at 89% for IC Team students; however, goal attainment was only calculable in 53% of all IC Team cases due to inconsistencies in documenting goals and data. In addition, this study suffered from small sample sizes of 17 IC Team cases total, and only 9 cases for which GAS was accurately documented. Low levels of goal documentation have also been cited in other program evaluation reports of IC Team, suggesting a need to better understand the processes that lead cases to either attain goals or fail in documenting the process such that goal attainment measures are not available (Gravois, 2000; Gravois & Rosenfield, 2001). In addition, while these goal attainment data provide preliminary support for positive IC Team impact on student outcomes, it is difficult to generalize the findings to all IC Teams cases when so many are documented incompletely.

Level of Implementation (LOI)

Level of implementation has been defined as the extent to which certain critical dimensions of a treatment program or initiative have been implemented as planned (Fudell, 1992). It has been recognized that the success of teaming models cannot be evaluated without considering the integrity with which the model is actually

operationalized in a given setting. Thus, the success or failure of any team model or intervention can not be attributed to that model unless it is certain that the crucial elements of the process have been implemented consistently and that alternative elements have not been implemented that might interfere with the outcome (Gresham, 1989; Nelson & Smith, 1991; Sheridan & Welch, 1996).

Despite recognition that implementation integrity is necessary for program evaluation studies, there has been a paucity of research citing the integrity of processes under which teaming models, and the intervention strategies they employ, are operating. Studies that are available lack consistency in the procedures and measures used to evaluate integrity. In addition, many integrity studies focus only on the level to which a single intervention is operationalized as opposed to measuring elements of a delivery system surrounding the intervention as a whole. Without information on implementation, evaluation of consultation models is unreliable and evaluation of their critical components is near impossible (Gresham & Gansle, 1993; Gresham & Kendell, 1987; Witt, Gresham & Noell, 1996).

Using level of implementation to evaluate teams. Implementation studies of teaming and consultation models have primarily evaluated whether differing levels of program implementation produce different outcomes desired by the program. In a review of school-based behavioral intervention studies, Gresham and Gansle (1993) found that for studies that measured intervention implementation, there was a moderate correlation between treatment integrity and treatment effect size. Thus, the more consistently interventions were implemented as planned, the greater the likelihood that desired treatment effects would result. Several educational models have cited implementation

integrity to be a crucial component in considering student outcomes related to the program.

Haynes, Emmons & Woodruff (1998) found that quality of implementation of a school development program correlated significantly with students' achievement in both mathematics and reading on standardized achievement tests across several schools, grades and geographic locations. Fuchs and Fuchs (1989) measured the impact of full versus partial implementation of Behavioral Consultation (BC) on student outcomes, by measuring outcomes at three differing levels of treatment implementation. The shortest condition included only two of the stages of BC, the middle condition included three of the stages of BC, and the final condition included all four BC stages. Decreased teacher perceptions of target behavior severity were evidenced for the two more intensive intervention groups when compared to the control group, and all intervention groups demonstrated decreased problem behaviors when observed in a classroom. However, this study also produced conflicting evidence in that the third complete condition of BC actually showed the lowest decrease in observed problem behavior, suggesting that more extensive conditions of implementation did not produce more beneficial outcomes.

Other studies argue that implementing programs with low levels of implementation are akin to providing no treatment. In the 1999 study by Kovaleski et al. of Instructional Support Teams (ISTs) in 117 elementary schools, it was demonstrated that students under conditions of high IST implementation made greater gains in the areas of reading comprehension, task completion and time on task than students under conditions of low IST implementation. Gains in comprehension were noted to occur when measuring students' performances from pre-test to post-test and from post-test to

follow-up, 45 days and 80 days after the initial identification of the student. In addition, while high implementation IST students' comprehension, task completion, and time on task continued to improve following IST intervention, low implementation IST students' behaviors actually declined during the post-intervention period. In many cases, low implementation conditions produced behavioral change indistinguishable from the change demonstrated by non-IST students.

Problems with measurement of level of implementation. While there is some consensus that implementation integrity and intensity leads to beneficial student outcomes, variations in measurement of integrity and outcomes have produced conflicting evidence. Whether implementation integrity contributes to student outcomes depends highly on the implementation and outcome measures used. In a study of Behavioral Consultation (BC) implemented in MATs, Fuchs, Fuchs, and Bahr (1990) created a short and long implementation condition in which length of time a student engaged in a self-monitoring intervention was varied. While both intervention groups benefited from increased goal attainment and improved behavior over time when compared to controls, the two implementation conditions did not differ in terms of teacher perceptions of specific behavioral problems, student or teacher satisfaction. Behavioral observations of student problem behaviors also did not reveal differences in problem behaviors between groups or in goal attainment rates. Although varying intervention length could be seen as a viable condition for differing treatment integrity, this study was limited in that it only varied the length of intervention by one day of student self-monitoring. In addition the use of subjective measures and varying definitions of effectiveness may have limited the generalizability of this study.

In another study, Wickstrom, Jones, LaFleur, and Witt (1998) demonstrated a 3-5% decrease in perceived and observed child disruptive behavior despite overall low levels of implementation integrity. Their analysis of 29 behavioral consultation cases revealed that teachers only responded to an average of 4% of student behaviors with a planned intervention consequence. Yet two measures of treatment outcomes indicated improvements in student behavior as teachers perceived student problems to decrease and recorded fewer observed episodes of problem behavior. Although observations of the intervention suggest extremely low integrity, a teacher-recorded implementation measure averaged at 54% integrity, and observations of teacher use of intervention stimuli (behavior charts, prompts etc.) averaged at 62%. Thus, while observation of intervention consequence use, observation of intervention stimuli use and teacher report measures may all seem viable measures of implementation integrity, these sources measure very different elements of the process, making it difficult to conclude whether any of the measures provide a true indicator of implementation that can be used to understand behavioral outcomes.

Robbins and Gutkin's 1994 study provides further indication of the extent to which implementation measures can vary. In their study of three behavioral consultation cases, teachers all reported that the intervention was implemented as planned; however, use of planned intervention methods only occurred 0-.5 times in a 30 minute period. This study was limited, however, by a very small sample size and the role of the consultant as interviewer for the self-report integrity measure, which might have influenced the consultee to provide less valid information. In addition, while positive verbalizations towards student behavior were observed as the direct measure of treatment integrity,

there were often two or three other interventions that the teacher agreed to use that were not part of the integrity measure. Thus, in reviewing the literature on treatment integrity, it is apparent that there are variations in measurement procedures. These studies also raise considerable concern that school-based teams suffer from such low levels of implementation integrity that measurement of outcomes may be unreliable. Variations in measures of treatment integrity and the prevalence of cases for which little implementation integrity is apparent suggest a need to come up with consistent measures of integrity known to be important in consultation and school-based interventions.

Evaluating teams through use of critical indicators. A few implementation integrity studies have measured the presence of certain critical components of a model in order to determine the necessity of each component to the effectiveness of the model. Specifically, research evaluating problem-solving teams and consultation identifies several elements of the process which seem to be important in determining successful outcomes. Some typically examined quality indicators of prereferral interventions have been identified as: a) behavioral definition of a target behavior, b) measurement of a baseline level of a behavior pre-intervention, c) design of a systematic intervention plan, d) implementation of the intervention as planned, e) graphing of intervention results, and f) comparison of baseline and post-intervention performance data (Flugum & Reschly, 1994; Reynolds & Gutkin, 1984).

Flugum and Reschly (1994) conducted a component analysis to measure the relation of each of these six indicators to student outcomes such as improved student behavior, degree of improvement, accomplishment of goals, and improved student functioning. The presence of indicators such as having a systematic intervention plan,

implementing the intervention with integrity and comparison of baseline and post-intervention data correlated with student behavior improvement. In addition, implementation integrity related to improved student functioning and graphing of data related to the degree of behavioral improvement that was observed. However, the responses of teachers and related service personnel in matched cases varied extensively, such that when examining teacher responses alone, few indicators seemed significantly related to outcomes. For both teachers and related service providers only the variable of behavioral improvement correlated with the number of quality indices observed. However, the presence of each of the quality indices in any given case was quite low, such that 56% of teachers and 65% of related service personnel responded that fewer than three of the six indicators were implemented. This study provides preliminary validity to the notion that certain components of prereferral intervention implementation may be related to the success of interventions as measured by various student outcomes. In addition, it suggests that low implementation of key prereferral indices may be the reason that past research on the effectiveness of PI teams has been inconsistent.

A number of other studies have documented the impact of certain characteristics or components of prereferral teams in producing desired student and team outcomes. Bergan & Tombari (1976) identified that consultant skill contributed significantly to determine whether problem identification steps took place, and that problem identification contributed to plan implementation which led to steps in problem solution. Although only 30% of cases reached the problem solution phase, cases reaching problem solution demonstrated student achievement of goals in 97% of the cases. Similarly, Graden, Casey & Bonstrom (1985) found that schools implementing successful PIs were

those supported by at least one individual for whom consultation was their primary role. Additional problem-solving steps were observed to be important for multi-disciplinary teams in a study by Telzrow et al. (2000). Factors relating to student goal attainment included definition of the problem in behavioral terms, collection of baseline data, identification of clear goals, collection of data during the intervention process, implementation of an intervention plan, and comparison of data with baseline. Most factors related only modestly to goal attainment; e.g., identification of clear goals and collection of data during the intervention accounted for 8% of student change. Finally, in a study of Teacher Support Teams, the use of appropriate forms and documentation throughout the process related to higher implementation of the process (Fudell, 1992).

Thus, it seems that there are several common components of the problem-solving process that when implemented consistently may lead to beneficial student outcomes in terms of goal-attainment and achievement. Specifically, it appears that defining the problem in clear terms, setting specific goals, implementing an agreed upon intervention, and consistently monitoring the intervention and data collection are crucial components leading to success of prereferral teams. However, as the literature varies in the components that are measured and relationships are modest, additional research is needed to support these findings (Bergan & Tombari, 1976; Flugum & Reschly, 1994; Levinsohn, 2000; Telzrow et al., 2000).

Level of implementation in IC Teams as measured by the LOI-R. Level of implementation in IC Teams has also been recognized as a crucial component in determining the success of the model. The Level of Implementation-Revised interview (LOI-R) conducted to evaluate IC Teams consists of four structured interviews

administered to the team, principal, teacher and consultant of each case. The interviews measure many of the quality indices cited in the literature including data collection at baseline through intervention evaluation stages, identification of clear goals, design and implementation of a systematic intervention, and comparison of pre and post-intervention data in evaluating the success of modification of an intervention. In addition, the LOI-R's include an agreement component in that both the consultee and the consultant have to agree on certain items regarding the intervention process and data in order to get full credit for implementation. Thus, while the LOI-R measures do not include direct observation of any intervention elements, structured items in which the consultant and teacher have to provide consistent information increases the likelihood that these components were actually present in the implementation.

Reliability and validity of the LOI-R. The LOI-R measure has been examined for inter-rater reliability of scoring as well as test-retest reliability. Fudell (1992) measured inter-rater reliability coefficients of .88 and .92 for two data collection periods. When individuals were administered the LOI-R measure on two occasions, their test-retest reliability was measured to be .78 and .88 for two data collection periods. Thus, the LOI-R scale demonstrates adequate inter-rater and test-retest reliability. McKenna (2005) assessed the validity of the LOI-R by comparing scores on the interview measure with actual consultation behaviors as measured by listening to audio tapes of the consultation sessions. The LOI-R dimensions were in agreement over 70% of the time with the actual taped communications. In addition, none of the process component scores differed significantly when measured using the LOI-R versus the tape scored measure. Thus, the

LOI-R was determined to be a valid measure of the seven consultation process components as they actually occur in consultant-consultee dyad communications.

Utility of the LOI-R. Preliminary studies of implementation integrity suggest that certain IC Team components are particularly important for student outcomes. In a study of the integrity of the IC Team process, collection and graphing of baseline and intervention data correlated positively with change in reading achievement scores measured through a criterion-referenced assessment of reading (Levinsohn, 2000). Implementation integrity of the IC Team model has been found to be highly variable between schools, however, so further research is needed both to delineate the conditions under which high and low implementation occurs and to provide insight into implementation conditions overall (Vail, 1996). In addition, research is needed to determine how implementation conditions relate to important outcome variables such as student goal attainment.

Summary

Since the passage of PL 94-142, school districts have used prereferral teaming models to improve instruction for difficult to teach students within the regular education environment. Recent laws such as No Child Left Behind (NCLB) and the Individuals with Disabilities Education Improvement Act (IDEIA) mandate that schools use evidenced based instructional practices. Thus, there is a need to critically evaluate the effectiveness of the various prereferral and teaming models in terms of student outcomes. In addition, to understand the context under which certain outcomes result, it is necessary to consider the integrity of the model as it is implemented in a given setting. IC Teams is an emerging prereferral model in which structuring and monitoring the collaborative

process is a main focus. While reduction of inappropriate referrals to special education is a consistently documented outcome of IC Teams, there is a need to understand what elements of the process, when implemented consistently, lead to beneficial student outcomes. Thus, the goal of this research is to examine the level of treatment integrity in IC Team cases and determine how implementation of the various components of the process relates to student goal attainment outcomes.

Chapter 3

Methodology

Purpose

The purpose of this research was to examine the implementation integrity of consultation cases and determine whether cases demonstrating a high degree of implementation of the IC Team Model result in differential student goal attainment from cases demonstrating lower levels of implementation. This research also explored how elements of the implementation process differ for cases that do not reach goal attainment and for those which goal attainment is unspecified or incalculable.

Cases were selected from archival data files based on their participation in the IC Team project and whether or not they had reached the intervention implementation stage of the process. Being in the implementation stage was necessary so that cases examined have initiated the process of change expected to influence goal attainment. Measures included the Level of Implementation (LOI-R) Interviews and the Student Documentation Form (SDF). Data were used from the SDF to produce goal attainment scores using Goal Attainment Scaling (GAS). Goal attainment was examined as an outcome measure because it allows for individualized measurement of outcomes for students in several grade levels and with various presenting concerns in either the academic or behavioral domain. This section describes the methodology used in this study.

Participants

The participants were 417 teacher-consultant units working within the IC Team project on an academic or behavioral concern of a student. Participants' data were

sampled from archival data collected routinely as part of the IC Team evaluation process. Units were drawn from 60 IC Team schools in 18 counties across the states of Maryland, Delaware and North Carolina. Data were drawn from schools involved in the IC Team project between 2000 and 2004.

The participant sample consisted of every case unit within the 60 schools in which a teacher requested assistance from the IC Team, was paired with a consultant, and the intervention implementation stage of the consultation process was reached. Because many of the schools were sampled for more than one project year and each team consisted of the same of group of professionals for several years, it is possible that there may be repeated sampling of the same consultants; however, each case represents a different consultant-teacher-student grouping.

Measures

Level of Implementation-Revised Interview (LOI-R). The LOI-R measures consist of four separate interviews designed to measure the extent to which the IC Team model has been operationalized for each case. Interviews were conducted with the teacher, consultant, principal and team by either graduate assistants working in the Lab for IC Teams or by IC Team facilitators from neighboring schools. Each interview consists of 5-20 scripted items that are read to the interviewee and scored using the LOI-R administration and scoring guide (Fudell, Gravois & Rosenfield, 1998). Items are scored as either a 1 or 0 depending on whether the respondent provided the correct answer dictated by the manual. For many items, it is also necessary for the consultant and teacher to provide answers demonstrating agreement about how the process was implemented. Each administrator of the LOI-R Interview was trained using the same

step-by-step module in which each item is read and explained, the documentation and scoring process is modeled, and opportunity to practice delivery of questions, documentation and scoring is provided. Samples of each of the scripted LOI-R interview measures are provided in Appendices A and B.

A LOI-R profile was compiled for each case that includes scores on three primary components of the IC Team model: collaborative process, delivery system, and overall level of implementation; and on fourteen individual subcomponents within these three areas. Seven of the subcomponents fall under the primary category of collaborative process and include:

1) Clear, accurate communication;
2) Contracting for the four elements of the collaborative relationship and obtaining agreement to work together;
3) Problem identification of the discrepancy between current and expected performance, and engagement in analysis of academic or behavioral concern;
4) Intervention recommendation of an agreed upon treatment plan designed to increase the desired performance;
5) Intervention implementation of the treatment plan while engaging in frequent graphing and monitoring of data;
6) Evaluation and follow up such that data are used to make decisions about effectiveness, and intervention is monitored as needed;
7) Curriculum Based Assessments that evaluate the target behavior in the classroom environment and use assessment for monitoring ongoing student progress.

The other seven subcomponents fall under the category of the delivery system and include: use of the referral process, referring teacher participation, systems management, organization and monitoring of cases, use of documentation and forms, team functioning and support team characteristics. For each of the seven collaborative process subcomponents a score was calculated based on the presence or absence of the key indicators of each component. For example, the contracting score was calculated from items one and two on both the teacher and case manager interviews (T1, T2, C1 and C2). If the dyad interviewed responded that they agreed willingly to work together for items T2 and C2, but did not describe that the critical four elements of the relationship were contracted for on items T1 and C1, then the score would be 2 out of 4, or 50% implementation of the contracting measure. In general, overall scores of 80% or greater are considered to have reached the criterion level of implementation (Vail, 1996).

Student Documentation Form (SDF) and Goal Attainment Scaling (GAS). The Student Documentation Form was developed to assist the consultative dyad in following the IC Team process and documenting case information. The SDF guides the collaborative process by prompting the dyad to follow necessary steps of the process. It also provides space for the dyad to document meeting information, student concerns, baseline performance, goals and outcome data. Scoring of SDFs was completed using the SDF Form Review (Appendix C) to provide a measure of the integrity with which the IC Team process was documented. The SDF Form Review scoring process also includes the calculation of a goal attainment score using Goal Attainment Scaling (GAS).

Goal Attainment Scaling was developed by Kiresuk et al. (1994) for the measurement of goals attained by individuals undergoing mental health interventions.

GAS has since been applied for several uses within the fields of medicine, public health, education and psychology. The GAS process in IC Teams involves setting specific goals for student performance in an identified area of concern and measuring the outcome of an intervention in terms of these goals. GAS is used both to guide the process and measure outcomes of IC Teams. Following the intervention, a goal attainment score is assigned based on the extent to which a student made progress in comparison to the baseline level of performance and in comparison to the goals that were set. To obtain a goal attainment score, the SDF is examined for three necessary pieces of information: 1) a statement of current student performance in observable/measurable terms (baseline), 2) a statement of a short-term, intermediate or long term goal, and 3) graphed data collected following an intervention. While other information is provided on the SDF, these three components alone are critical for Goal Attainment Scaling, and hereon will be referred to as "critical SDF components."

Goal Attainment Scaling is used by the consultative dyad to make decisions about the success of the intervention and to plan intervention modifications. Final student goal attainment is also described on the SDF Form Review as part of program evaluation and reflects the extent to which students obtained the goals set by the consultation dyad. Numerical values were then assigned to each of the descriptors to capture the leveled nature of the goal attainment descriptors. As noted in Gravois and Rosenfield (2001) and Levinsohn (2000), critical SDF components are often missing from documentation. For this reason, different scoring methods were used depending on the extent that critical SDF components were documented. On the SDF Form Review (Appendix C), these scoring methods are referred to as GAS A, GAS B and No GAS.

Cases were categorized according to the completeness of critical SDF component documentation. Cases that documented all of the critical SDF components necessary for goal attainment calculation are referred to as having "full documentation" and are scored according to GAS B on the Form Review. Possible scores obtained through the GAS B scoring method are (-1) data indicate trend below baseline, (0) data indicate trend consistent with baseline, (1) data indicate trend toward short term goal, (2) short term goal met, (3) interim goal met, and (4) long term goal met. For cases providing baseline performance and data following intervention, but no set goals, the GAS A scoring method is used and a score is given based on whether the progress data collected were (-1) below baseline data, (0) consistent with baseline, or (1) above baseline performance. These cases are referred to as having "partial documentation" of the critical SDF components. Cases that do not provide a statement of student baseline or data following intervention do not allow for goal attainment to be calculated, and are referred to as having "insufficient documentation" of critical SDF components. These cases represent those for which No GAS scoring can be completed. While these groups often do document one or two of the critical components, they are referred to as having "insufficient documentation" because the data missing are such that no determination about goal attainment or student progress can be made.

Procedure

A proposal was submitted to and approved by the Institutional Review Board at the University of Maryland to conduct this research using archival IC Teams data.

Data collection. Data from each case were collected as part of the IC Team end of year evaluation process during May and June following each of the four years of

project implementation. LOI-R interviews were conducted in person or over the phone by a graduate assistant at the Lab for IC Teams or by a facilitator from another school. All facilitators and graduate assistants were trained using a training module which included observation and explanation of the process of administering the LOI-R measure, scoring an interview from a practice case while receiving feedback, and practice administering an LOI-R for a real case situation with feedback. All LOI-R's were scored by the graduate assistant or facilitator conducting the interview. Goal attainment was scored from the SDF for each case according to the SDF Form Review- Revised (Gravois, 2001) by graduate assistants at the Lab for IC Teams or by facilitators from another school. Training for the goal attainment scoring included modeling and instruction of the scoring process and opportunity for practice and feedback with real case documentation. The scored measures and audio tapes were stored in a secure filing cabinet in the Lab for IC Teams and scores were entered into a computer database with identifying information removed.

Data analyses. The five research questions were investigated as follows:

Research Question 1: For IC Teams cases, how well are critical SDF components documented?

For this question, descriptive analyses were conducted to provide percentages of cases that provided full, partial and insufficient documentation.

Research Question 2: For cases that have not fully documented critical SDF components, what components are missing from the documentation?

This question was answered by examining the raw data to provide the frequency with which each of the critical SDF components was missing. Thus, the percent of cases lacking a baseline score, goals and data following intervention were reported.

Research Question 3: Do cases with full, partial and insufficient documentation of critical SDF components have different scores on the Level of Implementation- Revised Collaborative Process dimensions?

This analysis included reporting the treatment integrity levels obtained by all of the cases. In addition, treatment integrity levels were reported separately for cases with full, partial and insufficient critical SDF documentation. Next, a one-way analysis of variance was conducted between cases with full, partial and insufficient SDF documentation to determine if the seven dimensions of the collaborative process differed at the varying levels of documentation.

Research Question 4: For IC Team cases providing full documentation of critical SDF components, what are the levels of goal attainment?

Frequency analyses were conducted for each case providing full documentation of critical SDF components. Percentages of cases reaching long term, interim and short term goals were reported. In addition, frequencies of cases demonstrating progress toward a stated goal, consistent with baseline, or below baseline were reported.

Research Question 5: Do different levels of treatment integrity of the IC Team model, as measured by the Level of Implementation Collaborative Process scores, relate to different goal attainment rates, when critical components are fully documented?

This question examined the relationship between the Collaborative Process scores on the LOI-R and student goal attainment when components were fully documented. A

correlational analysis was conducted between the overall Collaborative Process score and the level of goal achieved by the student as measured by the GAS. Correlational analyses were also conducted for each of the seven process dimensions to determine which measures of implementation are related to goal attainment.

Chapter 4

Results

In this chapter, the results of the analyses are discussed for each of the presented research questions. For research question 1, the number of cases at each of the varying levels of documentation is reported. Analyses for research question 2 include descriptions of the type of data missing from incompletely documented cases as well as the frequency with which each type of data was missing. For research question 3, implementation levels for each type of case are described, and cases with varying levels of documentation are analyzed for differing implementation integrity. A description of levels of goal attainment for fully documented cases is provided for research question 4, and question 5 investigates any relationships that occurred between treatment integrity and levels of goal attainment.

Research Question 1: For IC Teams cases, how well are critical SDF components documented?

Descriptive analyses were conducted using the Student Documentation Form data to categorize cases based on the extent to which they documented critical SDF components. Cases were categorized as either having full documentation, partial documentation or insufficient documentation. Cases for which no goal attainment score could be determined were referred to as having "insufficient documentation" of critical SDF components. These cases either did not provide a student documentation form at all, or the form was lacking either baseline data points or data following intervention. "Partial documentation" of critical SDF components describes cases that documented baseline data points and data following intervention, but were missing documented goal

statements. Cases described as having "full documentation," provided all of the critical SDF components including at least one baseline data point, a short term/ interim/ or long term goal, and documented data following the intervention.

Of the 417 cases, 58.3% (n = 243) provided full documentation of critical SDF components. Thus, 41.7% of the sample of IC Team cases that reached intervention failed to document one or more of the critical SDF components. Of these cases, 33.8% (n = 141) provided insufficient documentation of the critical SDF components, including 11% (n = 45) which provided no student documentation form at all. Finally, cases with partial documentation of critical components comprised 7.9% (n = 33) of the sample.

Research Question 2: For cases that have not fully documented critical SDF components, what components are missing from the documentation?

In order to better understand which critical SDF components were missing when full documentation was not provided, additional descriptive analyses were completed for the items on the student documentation form measuring baseline data, goal setting, and data collected following the intervention. Descriptive analyses were conducted for missing components from cases with partial documentation, insufficient documentation and the two groups combined. The analyses revealed that 26.0% of the cases missing critical SDF documentation (45 out of the total 174 cases missing documentation) provided no documentation form at all. All of the cases providing no SDF form were categorized as having insufficient documentation. Cases missing a complete statement of student baseline performance constituted 31.6% of cases with insufficient or partial documentation. A measurable short term goal was not documented in 35.6% of cases lacking full documentation and data following the intervention were not documented in

28.2% of cases. Information on whether long term or interim goals were documented was not provided because this information is considered to be "optional" on the SDF Form Review and was not found consistently in archival files. A description of the frequency of cases missing critical documentation is presented in Table 1.

Table 1

Frequency of Cases Missing Essential Student Documentation Form Components in both Partial and Insufficient Documentation Groups

	Partial *(n = 33)*	Insufficient *(n = 141)*	Total *(N = 174)*
Entire Documentation Form	0% (*n* = 0)	31.9% (*n* = 45)	26.0% (*n* = 45)
Baseline	48.5% (*n* = 16)	27.7% (*n* = 39)	31.6% (*n* = 55)
Short Term Goal	63.6% (*n* = 21)	29.1% (*n* = 41)	35.6% (*n* = 62)
Data Following Intervention	33.3% (*n* = 11)	27.0% (*n* = 38)	28.2% (*n* = 49)

Research Question 3: Do cases with full, partial and insufficient documentation of critical SDF components have different scores on the Level of Implementation- Revised Collaborative Process dimensions?

Treatment integrity levels, as measured by the LOI interviews, were examined for each category of SDF documentation as well as for the entire sample. Treatment integrity of the total Collaborative Process could not be calculated for 23.7% (*n* = 99) of

the 417 cases due to the fact that data were missing from the archival files. For the seven process components of implementation, data were also missing. Thus, the means and standard deviations reported reflect the integrity scores from the maximum number of cases that could be calculated. Mean treatment integrity scores and sample sizes are reported in Table 2 for the Overall Collaborative Process dimension and each of the seven subcomponents.

A one-way analysis of variance was conducted between cases at the three differing levels of critical SDF documentation to determine if the seven dimensions of the collaborative process differed at the varying levels of documentation. Differences in mean integrity scores between groups were detected for the Overall Collaborative Process dimension and the Clear Communication, Contracting, Problem Identification, Intervention Recommendation, Intervention Implementation, and Intervention Evaluation subcomponents ($p < .001$ for all analyses). The critical values and significance levels are reported in Table 3.

Post hoc comparisons using the Tukey *a* test were conducted to reveal where differences occurred between groups. Post hoc procedures revealed that cases providing partial or insufficient documentation of critical SDF components had lower treatment integrity as measured by the LOI-R in the areas of Clear Communication ($p < .001$ for both), Problem Identification ($p < .001$ for both), Intervention Recommendation ($p = .038$, $p < .001$), Intervention Implementation ($p = .002$, $p < .001$), Intervention Evaluation ($p < .001$ for both) and in the Overall Collaborative Process score ($p < .001$ for both) than did cases providing full documentation.

Table 2

M and SD of Treatment Integrity Scores in Groups with Full, Partial and Insufficient Documentation of Critical SDF Components

	Overall Sample	Partial	Full	Insufficient
Overall Collaborative Process	81.74 *SD* = 16.43 (*n* = 318)	72.25 *SD* = 18.71 (*n* = 27)	86.56 *SD* = 11.19 (*n* = 182)	76.03 *SD* = 20.09 (*n* = 109)
Clear Communication	82.50 *SD* = 19.20 (*n* = 375)	70.35 *SD* = 23.88 (*n* = 30)	87.67 *SD* = 14.32 (*n* = 220)	76.29 *SD* = 22.23 (*n* = 125)
Contracting	87.56 *SD* = 22.53 (*n* = 414)	83.59 *SD* = 23.43 (*n* = 32)	89.98 *SD* = 19.86 (*n* = 242)	84.29 *SD* = 26.03 (*n* = 140)
Problem Id	85.88 *SD* = 20.12 (*n* = 378)	75.41 *SD* = 25.49 (*n* = 28)	90.95 *SD* = 14.29 (*n* = 222)	79.38 *SD* = 24.42 (*n* = 28)
Intervention Recommendation	84.98 *SD* = 27.23 (*n* = 404)	77.08 *SD* = 36.35 (*n* = 32)	89.50 *SD* = 21.65 (*n* = 235)	79.08 *SD* = 31.70 (*n* = 137)
Intervention Implementation	80.45 *SD* = 30.16 (*n* = 387)	69.03 *SD* = 33.80 (*n* = 31)	88.00 *SD* = 24.41 (*n* = 230)	70.29 *SD* = 34.25 (*n* = 136)
Intervention Evaluation	71.92 *SD* = 38.75 (*n* = 400)	54.17 *SD* = 44.60 (*n* = 32)	79.51 *SD* = 32.95 (*n* = 231)	63.26 *SD* = 43.21 (*n* = 137)
CBA	91.49 *SD* = 17.83 (*n* = 332)	92.59 *SD* = 16.72 (*n* = 27)	92.91 *SD* = 14.99 (*n* = 194)	88.74 *SD* = 22.04 (*n* = 111)

Table 3

F Values for Group Differences on Treatment Integrity Measures

	F	*p*
Overall Collaborative Process	21.358	<.001**
Clear Communication	22.954	<.001**
Contracting	3.408	.034*
Problem Identification	19.234	<.001**
Intervention Recommendation	8.081	<.001**
Intervention Implementation	18.662	<.001**
Intervention Evaluation	11.821	<.001**
Curriculum Based Assessment	2.00	.137

Additionally, the cases that provided insufficient documentation of SDF components demonstrated lower treatment integrity on the Contracting dimension than those providing full documentation ($p = .045$). The full documentation group was the only group with mean integrity scores consistently at or above 80, the criterion level considered to reflect full implementation of the IC Team process. Cases that provided partial documentation did not differ on any dimension from cases that provided insufficient documentation. Curriculum Based Assessment integrity did not differ across any of the three groups. As the CBA dimension primarily measures procedures related to the identification of the problem and not goal setting or monitoring, it may not be as sensitive to missing SDF components critical for goal attainment documentation. Table 4 lists the pair-wise comparisons that were significant.

Table 4

Differences between the Full, Partial and Insufficient Documentation Groups on Measured Integrity Dimensions

	Differences Detected	*p*
Overall Collaborative Process	Full > Partial	<.001**
	Full > Insufficient	<.001**
Clear Communication	Full > Partial	<.001**
	Full > Insufficient	<.001**
Contracting	Full > Insufficient	.045*
Problem Identification	Full > Partial	<.001**
	Full > Insufficient	<.001**
Intervention Recommendation	Full > Partial	.038*
	Full > Insufficient	.001**
Intervention Implementation	Full > Partial	.002**
	Full > Insufficient	<.001**
Intervention Evaluation	Full > Partial	.001**
	Full > Insufficient	<.001**
Curriculum Based Assessment		

Table 5

Percentage of Cases in the Full Documentation Group Demonstrating Progress Towards and Obtaining Goals

	GAS B Value	%	*n*
Progress Below Baseline	-1	0.4	1
Progress Consistent with Baseline	0	5.3	13
Progress Toward Goal	1	14.8	36
Short-term Goal Attained	2	42.4	103
Interim Goal Attained	3	17.7	43
Long-term Goal Attained	4	19.3	47

Note: The percentages were calculated from the total sample providing full documentation of critical components (N = 243).

Research Question 5: Do different levels of treatment integrity of the IC Team model, as measured by the Level of Implementation Collaborative Process scores, relate to different goal attainment rates, when cases are fully documented?

This question examined the relationship between the Collaborative Process scores on the LOI-R as they predicted varying levels of goal attainment. For cases that provided full critical SDF component documentation, a correlational analysis was conducted between the overall Collaborative Process score and the level of goal achieved by the student as measured by the GAS. Correlational analyses were also conducted for each of the seven process dimensions to determine which measures of implementation were

related to student goal attainment, based on a six-point scale (-1 to 4). The correlations are presented in Table 6.

Table 6

Correlations between Goal Attainment Level and LOI-R Components for Cases with Full Critical SDF Documentation

	r	*p*	*n*
Overall Collaborative Process	.126	.089	182
Clear Communication	.155	.022*	220
Contracting	-.010	.878	242
Problem Identification	.069	.306	222
Intervention Recommendation	.061	.355	235
Intervention Implementation	.129	.050*	230
Intervention Evaluation	.248	<.001**	231
Curriculum-Based Assessment	.095	.189	194

Note: Missing data in archival files resulted in different cell sizes for each calculation.
* Correlations were significant at the .05 level.
** Correlations were significant at the .01 level.

Treatment integrity in the area of Clear Communication was positively related to level of student goal attainment ($r = .139$, $p = .022$). Integrity in the areas of Intervention Implementation and Evaluation was also positively related to student goal attainment ($p = .050$, $p < .001$). While the relationship sizes were small, cases in which the teacher and case manager agreed upon the referral concern, process for problem-solving, the intervention to implement, and the outcomes for evaluation were more likely to have

students that reached higher goals. Cases that reached the implementation stage, documented progress regularly, and modified the intervention when it was unsuccessful were also slightly more likely to have students that obtained high levels of goal attainment.

Chapter 5

Discussion

One goal of prereferral problem-solving teams is to improve student academic and behavioral performance prior to consideration for special education services (Villa & Thousand, 2003). However, objectively reported student progress findings are infrequent, and need to be considered in relation to how well the problem-solving process was implemented (e.g., Gresham, 1989; Nelson & Smith, 1991; Sheridan & Welch, 1996; Welch et al., 1999). Thus, there is a need to document various measures of treatment integrity as they relate to objective student achievement and behavior improvement. This study examined the levels of implementation (LOI) of the IC Teams problem-solving model and their impact on student goal attainment. In addition, because goal attainment data have been absent in many IC Teams cases, the impact of critical SDF goal components on LOI was also examined.

Summary of Results

Results indicated that while implementation was assessed to be high in the majority of IC Teams cases, many were missing crucial pieces of SDF goal attainment documentation. Cases with full documentation of critical SDF components had higher implementation integrity in most areas when compared to those providing partial or insufficient critical SDF component documentation. IC Teams cases that did provide complete data were demonstrated to show high levels of goal attainment. In addition, goal attainment was slightly related to higher scores on the LOI-R dimensions of Clear Communication ($p = .022$), Intervention Implementation ($p = .050$), and Intervention Evaluation ($p < .001$). Thus, consistent implementation of the IC Teams model, critical

SDF documentation and progress monitoring may be important for student academic and behavioral achievement.

The Importance of Goal Setting and Documentation

While IC Teams research has consistently demonstrated high levels of implementation integrity as measured by the LOI-R, this study calls for further consideration of the goal setting and documentation process in evaluation of IC Teams' outcomes. High levels of missing SDF documentation are also consistently demonstrated in IC Team research, and represent an element of the process that is not adequately accounted for in the LOI-R (Gravois & Rosenfield, 2001; Levinsohn, 2000; Moniodis, 1996). This finding, together with past research that other consultation models have failed to produce cases that reach the final problem solution stage (Bergan & Tombari, 1976), raises concerns that cases may not be completing the entire problem-solving process. Missing SDF documentation and outcome data are also problematic because goal attainment rates cannot be determined for a large number of cases. Without goal attainment rates, the impact of IC Teams on student performance is unknown. Moreover, critical SDF documentation appears to be related to the overall implementation of the IC Team process. Thus, cases that fail to document goal data seem to be more likely to implement other elements of the process inconsistently.

IC Teams cases documenting some of the critical SDF components had no better implementation integrity than those providing insufficient or no SDF documentation altogether. This finding is similar to that of Kovaleski et al. (1999) who demonstrated that partial program implementation of ISTs was similar in outcome to no program implementation. However, the small amount of IC Teams cases in the partial

documentation category suggests that partial documentation is also infrequent. The bimodal distribution of this sample into cases providing full critical SDF documentation and those providing insufficient SDF documentation suggests that the goal setting and documentation process may be an all or nothing endeavor, in terms of critical components. Thus, when consultation dyads do not feel competent to follow the process completely, the goal setting process and related documentation may not be completed at all. While these results do not definitively describe whether it is the failure of the dyad to set and monitor goals that contributes to lower SDF documentation, it is likely that lack of staff competence plays a role in the completion of documentation. These findings argue for greater training and monitoring of professionals' skills in goal setting and progress monitoring. Additionally, the SDF form itself may need to be examined to ensure that the importance of critical elements is emphasized and documentation of these components is facilitated.

Process Components Impacting Goal Attainment

High implementation of the problem-solving process is always desirable; however, certain components may be more influential for determining student goal attainment than others. The finding that communication is important in pre-referral interventions is consistent with previous research that reports increased student performance when feedback is provided to teachers (Mortenson & Witt, 1998; Noell, Witt, Gilbertson, Ranier & Freeland, 1997; Witt, Noell, LaFleur & Mortenson, 1997). Thus, when communication between the consultant and the teacher is used to reinforce implementation procedures and discuss student progress, students may demonstrate higher academic achievement. Using performance feedback has also been documented to

relate to increased implementation of academic and behavioral interventions, which is likely a requisite condition for improved student performance (Jones, Wickstrom & Friman, 1997; Mortenson & Witt, 1998; Noell et al., 1997; Witt et al., 1997). Consistent implementation of an academic intervention is also intuitively and empirically linked to improved student performance (Bergan & Tombari, 1976; Gresham & Gansle, 1993; Telzrow et al., 2000). Thus, the current study replicates past findings that professional collaboration and intervention integrity are related to student outcomes.

The small but significant correlation ($r = .248$, $p < .001$) detected between integrity on the Intervention Evaluation dimension and goal attainment is also consistent with the literature on goal setting and monitoring. Several studies have documented the relationships between student goal setting, graphing, and progress feedback and increased academic achievement (e.g., Flugum & Reschly, 1994; Fuchs et al., 1990, Schunk & Swartz, 1993; Telzrow et al., 2000). Thus, the process of data collection and ongoing progress evaluation in consultation may be similar to student goal setting and monitoring in that it causes the consultant and consultee to continually assess the effectiveness of the intervention and make changes if necessary. When an intervention is found to be ineffective, it may be more likely that changes will be made that will in turn increase student performance. Likewise, when interventions are found to be successful, the consultant dyad may be reinforced to continue implementation.

The importance of the IC Teams process of objective, data-based evaluation of student progress is also consistent with past research. Wesson (1991) found that teacher groups that monitored intervention effectiveness through curriculum-based measurement had students with greater gains in reading than teacher groups that monitored

interventions without structured data collection and analysis. Thus, when data are collected and assessed objectively, teachers may be better able to modify interventions to increase student progress than when interventions are monitored through vague, subjective measures.

The findings that the Clear Communication, Intervention Implementation and Intervention Evaluation dimensions were slightly related to goal attainment support the collaborative nature of the IC Team process and argue for ongoing, objective progress-monitoring. Frequent meetings between the IC Team consultant and consultee, inherent in the IC Team process, may contribute to the dyad's ability to address questions, resolve problems in implementation, and discuss the intervention's effectiveness. However, the small size of these relationships is also consistent with previous research (Levinsohn, 2000), suggesting the need to examine other factors that may also be impacting student goal attainment.

Implications for Practice

As No Child Left Behind and IDEA legislation have demanded that school systems use evidence based instructional practices, it is important to utilize educational innovations that demonstrate effectiveness in terms of student outcomes. This study replicates previous findings that IC Teams students reach objectively measured academic and behavioral goals in a large percentage of cases (Gravois & Rosenfield, 2002). However, the finding that many IC Teams cases do not provide critical SDF documentation suggests that there is a need for increased training for consultees in documenting of baseline data collection, goal setting, and progress monitoring.

School systems utilizing prereferral problem-solving models should recognize the need to implement the model with integrity. Specifically, training and ongoing evaluation should occur in the areas of collaborative communication, intervention implementation and intervention evaluation, as these may be important for goal attainment. The process of intervention evaluation is likely useful in practical contexts because it alerts the consultative dyad to a student's progress, allowing changes to be made for increased achievement.

Directions for Future Research

As SDF documentation and Goal Attainment Scaling can be viewed both as indicators of process integrity as well as measures of program effectiveness, further research is needed to clarify the relationships between low documentation and process integrity, as well as outcomes. The current study suggests that low documentation may be related to lower scores on other measures of implementation integrity. However, additional research is needed to determine exactly which elements of the teaming process contribute to lower documentation. Additionally, it is unclear whether low documentation represents a failure of the consultative dyad to set goals and monitor progress or merely a failure to complete paperwork related to this process. Future research should examine other indicators to determine if goal setting is occurring when SDF components are not fully documented. Staff training and competence in using GAS techniques could be examined to determine if training impacts documentation and use of GAS. Finally, the usefulness of the SDF should also be examined to determine whether the form facilitates goal setting and documentation in consultation sessions.

Because GAS and SDF documentation are often used for program evaluation, low documentation poses a problem in that effectiveness cannot be determined for a large proportion of cases. Because 42% of cases did not document complete SDF data, student progress cannot be measured in these cases. Thus, while the GAS method has been demonstrated to be useful for monitoring progress in a variety of situations, it may not be the most effective evaluation tool due to its reliance on participant reporting. Future research should investigate other program evaluation measures for evidence of student goal attainment. In addition, there is a need to clarify the relationship between SDF documentation and goal attainment. As 79% of cases with full documentation achieved goals, further analyses are needed to determine whether documentation procedures assist in student goal attainment. While this study could not determine goal attainment for cases missing essential documentation components, future research could examine other achievement measures to determine if consultation documentation impacts student performance on these measures.

Finally, additional analyses are needed to examine the small relationships between LOI-R process components and student goal attainment. While LOI dimensions may be important factors related to goal attainment, research is mounting to suggest that they alone may not be sufficient to determine whether goal attainment will occur (Levinsohn, 2000). Future models may be needed that account for student, school and professional variables in determining student achievement. Because goal documentation is so inconsistent, professional competence in goal setting and monitoring may be examined for its impact on goal documentation and subsequent attainment. The current study suggests that implementation integrity, SDF documentation, and goal attainment

are related but not equal processes. Thus, further analyses are needed to determine the other factors that impact implementation, documentation and achievement outcomes.

Limitations

This research is limited by its archival nature in that little can be determined about cases with missing information from the archival files. Using only archival files, it is impossible to tell whether SDF documents were not completed for a number of cases, or whether they were just not returned for program evaluation. Without this information, it is difficult to determine whether the data present is representative of all IC Teams cases, or just those who returned documentation. In addition, the archival files were also missing individual pieces of LOI-R and SDF data, thus limiting the number of cases on which some analyses could be done. Although one research question attempts to investigate the problematic nature of missing documentation, this factor is also inherently a limitation of the study as it resulted in a large number of cases for which goal attainment could not be determined. Another limitation of archival data was that additional analyses could not be conducted with participants to obtain explanations for study findings such as limited form completion and small *n* sizes in Group A.

The ability of this study to detect relationships between program implementation dimensions and goal attainment was also impacted by the amount of missing documentation and the narrow range of LOI-R scores. Because critical SDF components were only documented by cases that obtained goals, little can be inferred about cases that did not obtain goals. The model may also have better accounted for student goal attainment if other factors such as student, school and staff variables could have been controlled. The archival files did not provide sufficient matched information for these

variables, however. Additionally, the fact that high LOI scores were obtained by all of the cases limits the ability of correlation analyses to detect possible relationships. The small cell size both in the partial documentation group and in the cases not demonstrating progress towards goals may have impacted the ability of analyses to draw conclusions about these groups. Finally, as study period spanned four years, there may be an impact of time on implementation levels and goal attainment that is not accounted for by these data.

Summary

A main purpose of problem-solving teams is to assist students in meeting academic standards without the need for special education services (Villa & Thousand, 2003). As recent literature and legislation cite the need for models that increase objective student achievement, it is necessary to measure the components of various innovations to determine their efficacy in producing desired student outcomes. Additionally, innovations cannot be evaluated without knowing the extent to which they are implemented (e.g., Gresham, 1989; Nelson & Smith, 1991; Sheridan & Welch, 1996; Welch et al., 1999).

The results of this large scale, multi-year study replicate previous findings that describe high implementation integrity and high goal attainment levels in Instructional Consultation Team cases. However, they also replicate findings that IC Team cases are often lacking in critical SDF documentation of goal attainment. Without consistent form documentation, goal attainment rates must be viewed cautiously. The importance of SDF documentation for student achievement is not clearly understood. However, critical documentation of SDF components appears to be related to implementation integrity.

Consultation dyads that fail to document SDF components may not be engaging in critical elements of the IC Teams model such as goal setting and progress monitoring. Thus, professional training and competence should be examined to determine whether effective training is impacting use of the GAS and documentation process.

The ability of consultative dyads to communicate clearly, implement interventions with integrity and continually monitor progress data was related somewhat to student goal attainment. Thus, training programs may benefit from emphasizing collaboration, monitoring of intervention integrity and use of objective data. While the relationship of IC Team process dimensions to student goal attainment was small, a number of limitations may have impacted this finding. Future research should examine the role of training and professional skill in the development of goal setting behaviors, documentation of progress monitoring data, and goal attainment. Additional measures of student achievement may be useful to detect the impact of implementation integrity when other documentation is unavailable.

Appendices

Appendix A: Level of Implementation-Revised Case Manager Interview Form

	Process	Delivery		COMMENTS
C1	**Y N** (Tr1)		At your first meeting, how did you explain the problem-solving process to _____? (2a) _____ Consultation stages _____ Meaning of Collaboration _____ Time to meet _____ Parameters of confidentiality _____ _____	
C2	**Y N** (Tr2)		Did _____ agree willingly to work with the IC Team? (2b) YES NO	
C3	**Y N** (Tr3)		Describe the initial referral concern. What concerns did you and the teacher focus upon? What was the current/ baseline performance and goals established for the concern(s)? (3a) _____ _____	
			What activities did you and _____ undertake to identify the presenting problem?:(Check the activities described by Case Manager to identify the academic or behavioral problem-VERBAL DESCRIPTIONS)	
C4	**Y N**		**ACADEMIC (3b)** ___ Analysis of entry level skills using CBA/Inst. Assessment	
C5	**Y N**		___ Analysis of targeted academic task(s).(i.e.TASK OR ERROR ANALYSIS) Specify How? _____	
C6	**Y N**		_____ ___ Specification of terminal goal. What? _____ _____	
			BEHAVIOR (3c):	
C7	**Y N**		___ Possibility of academic problem assessed (3b).	
C4*	**Y N**		___ Analysis of entry level skills using CBA/Inst. Assessment in main academic areas or during time in which behavior occurs.	
C8	**Y N**		___ Analysis of setting and situation. How? _____	
C9	**Y N**		___ Analysis of antecedents/ consequences. How? _____	
C10	**Y N**		___ Specification of desired behavior. How? _____	

	Process	Delivery		COMMENTS
			What strategies or interventions did you agree to implement? Describe them. Who was responsible for each aspect? When was the intervention to take place? Strategy _____ Who? _____ When? _____ Strategy _____ Who? _____ When? _____	
C 11	**Y N** (Tr9)		There is agreement between Case Manager and Teacher as to which interventions to implement and strategy relates to identified concern? (4b)	
C 12	**Y N** (Tr10)		There is evidence of specification of who is responsible for what, when in intervention development? (4c)	
C 13	**Y N** (Tr11)		How was the effectiveness of the strategy/ intervention to be monitored? (4d) _____ _____	
C 14	**Y N** (Tr12)		Did you and the teacher meet to determine whether the intervention/ strategy was implemented as planned? YES NO Did you and the teacher agree as to how much modification was needed, if any? (5a) YES NO	
C 15	**Y N** (Tr14)		How many times did you have scheduled/ formal meetings with the teacher to discuss case progress? (5b)	
C 16	**Y N** (Tr16)		How was the decision to modify, continue, or terminate the intervention made? (6b) YES if based upon data/information NO if not based upon data/ information	
C 17		**Y N** (Tr18)	Did _____ participate in all meetings (including IC Team meetings) during which the referral problem was discussed, that is beyond brief updates? (10a) YES NO	
C 18		**Y N**	Did the teacher actively plan and make the decision as to which intervention to implement? (10b) YES NO	
C 19		**Y N**	How much time passed between the teacher's request for assistance (date of referral) and your first meeting? (12a) _____	
C 20		**Y N**	Do you or (teacher) have data generated from this case? (12b) YES NO	

Appendix B: Level of Implementation-Revised Teacher Interview Form

	Process	Delivery		COMMENTS
T1	**Y N** (C1)		What was your understanding of what the IC Team (collaborative problem-solving) process would be after your first meeting with the Case Manager? (2a) _____ _____	
T2	**Y N** (C2)		Did you agree to work on [the student's] problem with the Case Manager and Team? (2b) YES NO	
T3	**Y N** (C3)		Describe the initial referral concern. What concerns did you and the case manager focus upon? What was the current/ baseline performance and goals established for the concern(s)? (3a) _____ _____ _____ _____	
T4	**Y N**		What are some activities that you and the case manager undertook to better define the problem? **ACADEMIC** (3b) ___ Assessment of student's academic skills and instructional level. (7)	
T5	**Y N**		___ Assessments conducted in classroom material and is focused upon the individual student rather than norm group. (7b; 7c) **BEHAVIOR** (3c)	
T6	**Y N**		___ Assessment of student's academic skills and instructional level relevant to times/ situations of behavioral concern. (3a)	
T7	**Y N**		___ Analysis of antecedents/ consequences	
T8	**Y N**		___ Analysis of settings and situations	
			What strategies or interventions did you agree to implement? Describe them. Who was responsible for each aspect? When was the intervention to take place? Strategy _____ Who? _____ When? _____ Strategy _____ Who? _____ When? _____	

	Process	Delivery		COMMENTS
T9	**Y N** (C11)		There is agreement between Case Manager and Teacher as to which interventions to implement and strategy relates to identified concern? (4b)	
T10	**Y N** (C12)		There is evidence of specification of who is responsible for what, when in intervention development? (4c)	
T11	**Y N** (C13)		How was the effectiveness of the strategy/ intervention to be monitored? (4d) _____ _____	
T12	**Y N** (C14)		Did you and the case manager meet to determine whether the intervention was implemented as planned? YES NO Did you and the case manager agree as to how much modification was needed, if any? (5a) YES NO	
T13	**Y N**		Describe what type of information was collected during the intervention. How often was the information collected (6a)? _____ Was information graphed/ charted? (5c) _____	
T14	**Y N** (C15)		Did you have scheduled meetings with the case manager to discuss the student's progress? (5b) YES NO	
T15	**Y N**		After participating in the IC process for this case, how would you rate the outcome (listen to all choices, then decide): "We achieved.... ___ much more than expected ___ somewhat more than expected ___ what was expected, ___ somewhat less than expected ___ much less than expected (How do you know? (6a) _____ _____ _____	
T16	**Y N** (C16)		How was the decision to continue, modify, or terminate the intervention made? (6b) YES if based upon analysis of data NO if not based upon data	
T17		**Y N** (Tm4)	What did you do with the completed referral/ request for assistance form? (9b; 11b)	
T18		**Y N**	Did you feel that you were a contributing part of the problem solving process? _____ To the IC Team (10a)? _____ That your input was valuable? (10b) _____	

Appendix C: Student Documentation Form Review with Goal Attainment Scaling

System's Tracking Information:

SDF Review Part 1 should be completed for each case listed on the IC-Team System's Tracking Form. SDF Review Part 1 should be completed even if there is no SDF available to review for that case. The following information should be completed using the system's tracking form. A current and up-to-date copy of the System's Tracking Form should be obtained from the system's manager or team facilitator.

The following information should be completed using the system's tracking form. A current and up-to-date copy of the System's Tracking Form should be obtained from the system's manager or team facilitator.	Case manager's Name: _____ Teacher's Name: _____ Student Name: _____ School Name: _____ Request/Referral Date: _____ Date of 1st Contact/ Contracting: _____ Final Stage of problem-solving indicated: _____ Last date entry was made on tracking form: _____ Indication whether case was closed: _____ Reason for case closure: _____ COMMENTS:

General SDF Scoring:

SDF Review Part 1 should be completed for each case listed on the IC-Team System's Tracking Form. **Item 1** of SDF Review Part 1 should be completed even if there is no SDF available to review for that case.

Indicate whether the teacher or case manager has a SDF available for the case. Place a "+" if available or a "-" if not available.	SDF Available _____	**Item 1 Score** **Y** (if +) **N** (if -)
Review the 1st Page of the SDF for the presence of the following information. Place a "+" if the information is present or a "-" if not present.	Case manager's Name: _____ Teacher's Name: _____ Student Name: _____ Date Started: _____ COMMENTS:	**Item 2 Score** **Y** (if all +) **N** (if any -)
Review SDF consultation summary on back of form. Indicate whether the information is present or absent. "+" if present or "-" if not present.	SDF Consultation Summary: Date of each contact; _____ Brief summary of consultation: _____ Follow-up meetings and tasks: _____	**Item 3 Score** **Y** (if all +) **N** (if any -)

See SDF Review Form 2 to determine whether a SDF Concern can be scored.

Selecting the SDF Concern Area to Score:
An **SDF Part 2** should be completed for any case that has an SDF **and** has a concern area that has completed at least Step 3 of the Problem Analysis Stage (see Steps 1-6 on the SDF). The first concern to be scored should be that which has progressed furthest through the problem-solving stages as indicated by documentation on the SDF. Note: There may be up to four Part 2s per SDF (one for each concern that has progressed through the problem-solving stages enough to be scored).

Write the general description of the Concern to be reviewed.	Concern area to be scored:	
Review the GAS (on page 1 of the SDF) for the presence ("+")or absence ("-")of the following information for the concern being scored **NOTE: Optional information should be recorded but not included within the scoring of the item.**	A) General Statement of Concern _____ B) Instructional level considered _____ *(Is Y or N circled?)* C) Statement of current performance following baseline _____ *(Statement in words or numbers, as long as it is clearly a statement of baseline)* D) Measurable short-term goal with time specified _____ *Optional: Are there Interim / long-term goals with time specified? (_____)* COMMENTS:	**Item 4A Score** **Y** **N** **Item 4B Score** **Y** **N** **Item 4C Score** **Y** **N** **Item 4D Score** **Y** **N**
Review the Operational Definition of the concern (page 2 of SDF). Indicate whether the information is present ("+") or absent ("-"). **NOTE: Optional information should be recorded but not included within the scoring of the item.**	**Operational Definition:** What specific behavior will be recorded: _____ When will the behavior be recorded: _____ *Optional: Where will the behavior be recorded* (_____) COMMENTS:	**Item 5 Score** **Y** (if all +) **N** (if any -)
Review the graph for the concern (page 2 of SDF). Indicate the type of baseline information recorded. Place a "+" next to the best description of baseline recorded.	**Baseline for 1st Concern:** 0 baseline points recorded: _____ 1 baseline point recorded: _____ 2 baseline points recorded: _____ 3 or more baseline points recorded: _____ COMMENTS:	**Item 6 Score** **Y** (if 2 or more baseline points recorded) **N** (if less than 2 baseline points recorded)

SDF Part 2 Continues on back

Review the graph for the concern (page 2 of SDF). Record whether the indicated information is present ("+") or absent ("-"). **NOTE: Score "+" if unlabeled axis could reasonably be identified from information within the operational definition.**	**Graph Labels:** Clearly marked vertical axis _____ Clearly marked horizontal axis: _____ COMMENTS:	**Item 7 Score** **Y** (if all +) **N** (if any -)
STOP Stop scoring here if case has not begun Intervention Design		
Review the intervention description for the concern (page 2 of SDF). Record whether the indicated information is present ("+") or absent ("-"). **NOTE: Score "+" if intervention described on accompanying documentation.**	**Intervention Description:** What of intervention _____ When of intervention: _____ Who of Intervention: _____ COMMENTS:	**Item 8 Score** **Y** (if all +) **N** (if any -)
STOP Stop scoring here if case has not begun Intervention Implementation		
Review the graph, intervention description and consultation summary to determine if there is an indication ("+") of intervention implementation.	**Intervention Implementation:** Intervention implementation indicated on graph or on SDF Consultation Summary. _____ COMMENTS:	**Item 9 Score** **Y** (if +) **N** (if -)
STOP Stop scoring here if case has not begun Intervention Evaluation		
Review the graph, intervention description and consultation summary to determine if an evaluation of the intervention's effectiveness has been conducted ("+"). Intervention should be evaluated within 6 weeks of implementation. **NOTE: if data indicates progress toward stated goal, no indication of evaluation required. However, if progress is limited or declining, review SDF for indication of change in intervention.**	**Intervention Evaluation:** Data (beyond baseline) made weekly or regular basis with rationale provided: _____ Intervention Evaluation indicated: _____ COMMENTS:	**Item 10 Score** **Y** (if all +) **N** (if any -)
See SDF Review Part 3 to determine whether Goal Attainment can be scored.		

Goal Attainment for the SDF Concern Area:
Goal Attainment Scaling (**Part 3**) should be used for concerns in which SDF Review Part 2 has been completed and there are more than two weeks of data (or two data points) present following intervention implementation. Review the following Decision Boxes and determine which best describes the characteristics for the concern reviewed. Use the Goal Attainment Scale indicated in that Decision Box.

Decision Box #1:
- 0 baseline indicated
- no short-term goal
- no data collected following intervention implementation

No Goal Attainment

Decision Box #2:
- 1-3 baseline points indicated
- no short-term, no interim/ long-term goals
- data collected following intervention implementation

Use Goal Attainment Scale A

Decision Box #3:
- 1-3 baseline points indicated
- short-term or interim/ long-term goal indicated
- data collected following intervention implementation

Use Goal Attainment Scale B

Goal Attainment Scales:
Review the goal attainment scale selected. For Scale A, indicate ("+") the best description of goal attainment for the concern reviewed. For Scale B, indicate ("+") the highest or lowest level of goal attainment achieved. For example, if the short-term and interim goals were both met, place a "+" next to the interim goal.

Goal Attainment Scale A

_____ Progress above the baseline
_____ Progress below the baseline
_____ Progress consistent with baseline

Goal Attainment Scale B

_____ Long-term goal met (+4)
_____ Interim-goal met (+3)
_____ Short-term goal met (+2)
_____ Trend toward goal (+1)
_____ Trend consistent with baseline (0)
_____ Trend below baseline (-1)

Selecting additional concerns to review:
If the SDF being reviewed has additional concern areas that were addressed within the case, repeat the review process for the next concern using **SDF Parts 2 and 3**.

Staple or clip SDF Review Forms from the same case together!

References

Bergan, J. R., & Tombari, M. L. (1976). Consultant skill and efficiency and the implementation and outcomes of consultation. *Journal of School Psychology, 14*, 3-13.

Chalfant, J. C., & Pysh, V. D. (1989). Teacher Assistance Teams: Five descriptive studies on 96 teams. *Remedial and Special Education, 10*(6), 49-58.

Costas, L. T. (2003). Impact of instructional consultation on teacher satisfaction, skill development and skill application. Unpublished master's thesis, University of Maryland, College Park.

Flugum, K. R., & Reschly, D. L. (1994). Prereferral interventions: Quality indices and outcomes. *Journal of School Psychology, 32*, 1-14.

Fuchs, D., Fuchs, L. S., & Bahr, M. W. (1990). Mainstream Assistance Teams: A scientific basis for the art of consultation. *Exceptional Children, 57*, 128-139.

Fuchs, D., Fuchs, L., Gilman, S., Reeder, P., Bahr, M., Fernstrom, P., et al. (1990). Prereferral intervention through teacher consultation: Mainstream Assistance Teams. *Academic Therapy, 25*(3), 263-276.

Fudell, R. (1992). Level of implementation of teacher support teams and teachers' attitudes toward special needs students. Dissertation Abstracts International, 53 (05), 1399A. (University Microfilms No. AAC-9227463).

Fudell, R., Gravois, T. A., & Rosenfield, S. (1998). *Level of Implementation Interview-Revised*. Unpublished evaluation materials. The Laboratory for Instructional Consultation Teams, University of Maryland, College Park.

Glover, S., Burns, J., & Stanley, B. (1994). Goal Attainment Scaling as a method of

monitoring the progress of people with severe learning disabilities. *The British Journal of Learning Disabilities, 22*, 148-150.

Graden, J. L., Casey, A., & Bonstrom, O. (1985). Implementing a prereferral intervention system: Part II. The data. *Exceptional Children, 51*, 487-496.

Graden, J. L., Casey, A., & Christenson, S. L. (1985). Implementing a prereferral intervention system: Part I. the model. *Exceptional Children, 51*, 377-384.

Gravois, T. A. (2001). *The Student Documentation Form Review- Revised*. Unpublished scoring manual. The Laboratory for Instructional Consultation Teams, University of Maryland, College Park.

Gravois, T. A., & Rosenfield, S. (2001). Program Evaluation of the Instructional Consultation Team Consortium. Unpublished manuscript. The Laboratory for Instructional Consultation Teams, University of Maryland, College Park.

Gravois, T. A., & Rosenfield, S. (2002). A multi-dimensional framework for evaluation of Instructional Consultation Teams. *Journal of Applied School Psychology, 19*(1), 5-29.

Gresham, F. M. (1989). Assessment of treatment integrity in school consultation and prereferral intervention. *School Psychology Review, 18*, 37-50.

Gresham, F. M., & Gansle, K. A. (1993). Treatment integrity of school-based behavioral intervention studies: 1980-1990. *School Psychology Review, 22*, 254-273.

Gresham, F. M., & Kendell, G. K. (1987). School consultation research: Methodological critique and future research directions. *School Psychology Review, 16*, 306-316.

Haynes, N. M., Emmons, C. L., & Woodruff, D. W. (1998). School development

program effects: Linking implementation to outcomes. *Journal of Education for Students Placed at Risk, 3*, 71-85.

Jones, K. M., Wickstrom, K. F., & Friman, P. C. (1997). The effects of observational feedback on treatment integrity in school-based behavioral consultation. *School Psychology Quarterly, 12*, 316-326.

Kiresuk, T. K., Smith, A., & Cardillo, J. E. (1994). *Goal Attainment Scaling: Applications, Theory and Measurement*. Hillsdale, NJ: Erlbaum.

Kovaleski, J. K., Gickling, E. E., Morrow, H., & Swank, P. R. (1999). High versus low implementation of Instructional Support Teams: A case for maintaining program fidelity. *Remedial and Special Education, 20*, 170-183.

Kratchowill, T. R., Elliott, S. N., & Bussee, R. T. (1995). Behavior Consultation: A five-year evaluation of consultant and client outcomes. *School Psychology Quarterly, 10*, 87-117.

Levinsohn, M. R. (2000). Evaluating Instructional Consultation Teams for student reading achievement and special education outcomes. Unpublished doctoral dissertation, University of Maryland College Park.

Maher, C. A. (1983). Goal Attainment Scaling: A method for evaluating special education services. *Exceptional Children, 49*, 529-536.

Mortenson, B. P. & Witt, J. C. (1998). The use of weekly performance feedback to increase teacher implementation of a prereferral academic intervention. *School Psychology Review, 27*, 613-628.

Nelson, J. R., & Smith, D. J. (1991). Prereferral intervention: A review of the research. *Education & Treatment of Children, 14*, 243-254.

Noell, G. H., Witt, J. C., Gilbertson, D. N., Ranier, D. D., & Freeland, J. T. (1997). Increasing teacher intervention implementation in general education settings through consultation and performance feedback. *School Psychology Quarterly, 12*, 77-88.

Oren, T., & Ogletree, B. T. (2000). Program evaluation in classrooms for students with autism: Student outcomes and program processes. *Focus on Autism and Other Developmental Disabilities, 15*, 170-175.

Robbins, J. R., & Gutkin, T. B. (1994). Consultee and client remedial and preventative outcomes following consultation: Some mixed empirical results and directions for future researchers. *Journal of Educational and Psychological Consultation, 5*, 149-167.

Rosenfield, S. (1987). *Instructional Consultation*. Hillsdale, NJ: Erlbaum.

Rosenfield, S., & Gravois, T. A. (1996). *Instructional Consultation Teams: Collaborating for Change*. New York, NY: Guilford Press.

Rosenfield, S., & Gravois, T. A. (1999) Working with teams in the school. In *The Handbook of School Psychology* (pp.1025-1040). New York, NY: John Wiley & Sons.

Schunk, D. H., & Schwartz, C. W. (1993). Goals and progress feedback: effects on self-efficacy and writing achievement. *Contemporary Educational Psychology, 18*, 337-354.

Sheridan, S. M., Eagle, J. W., Cowan, R. J., & Mickelson, W. (2001). The effects of Conjoint Behavioral Consultation: Results of a 4-year investigation. *Journal of School Psychology, 39*, 361-385.

Sheridan, S., & Welch, M. (1996). Is consultation effective? A review of outcome research. *Remedial and Special Education, 17*, 341-355.

Sindaler, P. T., Griffin, C. C., Smith, S. W., & Watanabe, A. K. (1992). Prereferral intervention: Encouraging notes on preliminary findings. *The Elementary School Journal, 92*, 245-259.

Telzrow, C. F., McNamara, K., & Hollinger, C. L. (2000). Fidelity of problem-solving implementation and relationship to student performance. *School Psychology Review, 29*, 443-461.

Thousand, J. S., & Villa, R. A. (1992). Collaborative teams: A powerful tool in school restructuring. In R. A. Villa, J. S. Thousand, W. Stainback, & S. Stainback (Eds.), *Restructuring for Caring & Effective Education* (pp. 73-108). Baltimore, MD: Paul H. Brookes.

Vail, P.L. (1996). Instructional Consultation Teams: Analysis of Level of Implementation over two years and its relationship with team collaboration. Unpublished Master's thesis, University of Maryland College Park.

Villa, R. A., & Thousand, J. S. (2003). Making inclusive education work. *Educational Leadership, 61*(2), 19-24.

Weiner, R. K. (2002). A comparison of Instructional Consultation Teams to Individualized Education Plan Teams on special education referral and eligibility rates of minority students. Unpublished master's thesis, University of Maryland College Park.

Welch, M., Brownell, K., & Sheridan, S. M. (1999). What's the score and game plan on teaming in schools? *Remedial and Special Education, 20*, 36-49.

Wehmeyer, M. L., Hughs, C., Agran, M., Garner, N., & Yeager, D. (2003). Student-directed learning strategies to promote the progress of students with intellectual disability in inclusive classrooms. *International Journal of Inclusive Education, 7*, 415-428.

Wesson, C. L. (1991). Curriculum-based measurement and two models of follow-up consultation. *Exceptional Children, 57,* 246-256.

Wickstrom, K. F., Jones, K. M., LaFleur, L. H., & Witt, J. C. (1998). An analysis of treatment integrity in school-based Behavioral Consultation. *School Psychology Quarterly, 13*, 141-154.

Witt, J. C., Noell, G. H., LaFleur, L. H., & Mortenson, B. P. (1997). Teacher use of interventions in general education settings: measurement and analysis of the independent variable. *Journal of Applied Behavior Analysis, 30*, 693-696.